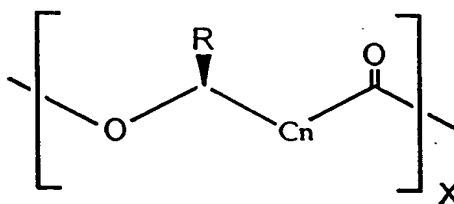


<u>R-group</u>	<u>Monomer</u>	<u>Abbreviation</u>
methyl	3-hydroxybutyrate	(3HB)
ethyl	3-hydroxyvalerate	(3HV)
propyl	3-hydroxycaproate	(3HC)
butyl	3-hydroxyheptanoate	(3HH)
pentyl	3-hydroxyoctanoate	(3HO)
hexyl	3-hydroxynonanoate	(3HN)
heptyl	3-hydroxydecanoate	(3HD)
octyl	3-hydroxyundecanoate	(3HUD)
nonyl	3-hydroxydodecanoate	(3HDD)



n = 1	3-hydroxyacyl monomer
n = 2	4-hydroxyacyl monomer
n = 3	5-hydroxyacyl monomer

FIG. 2

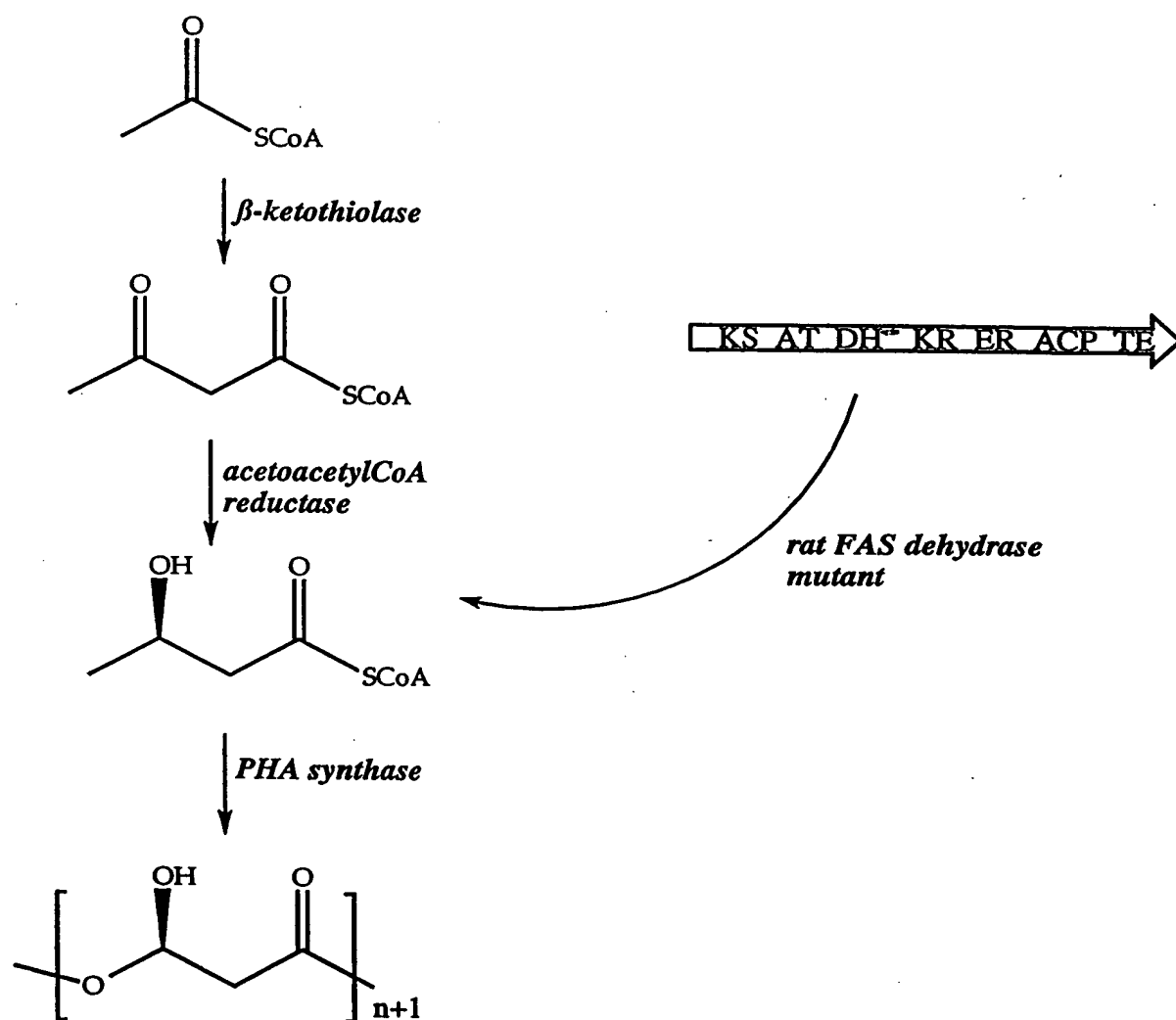


FIG. 3

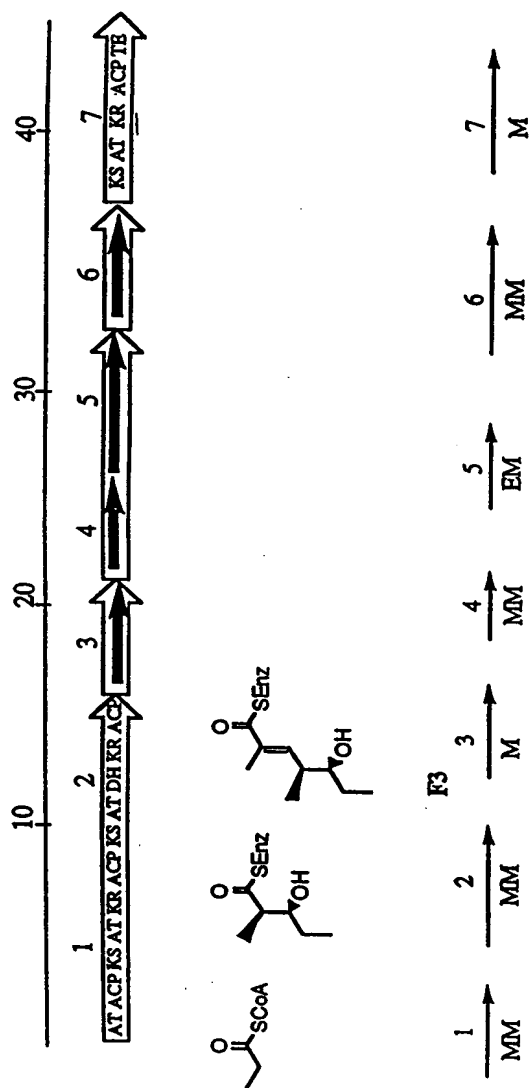


FIG. 4

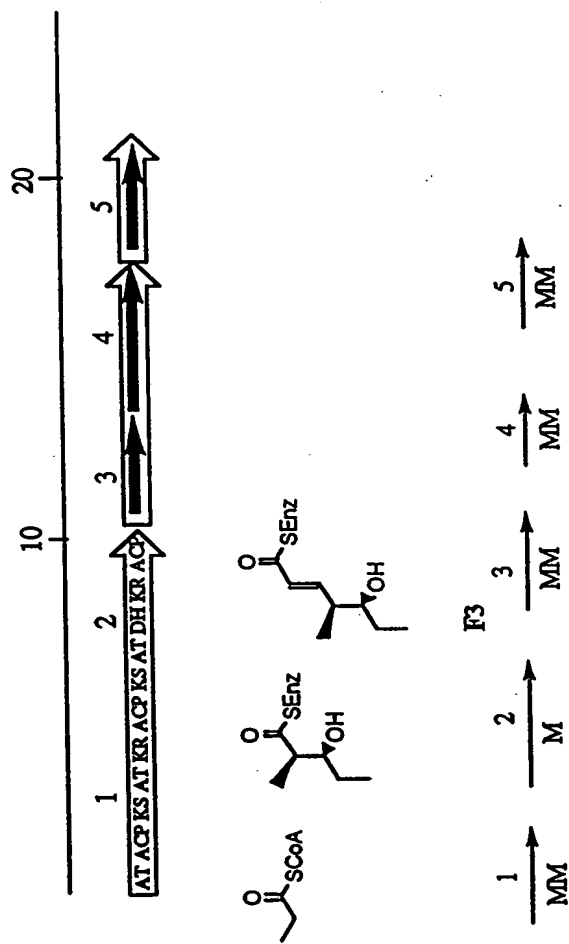
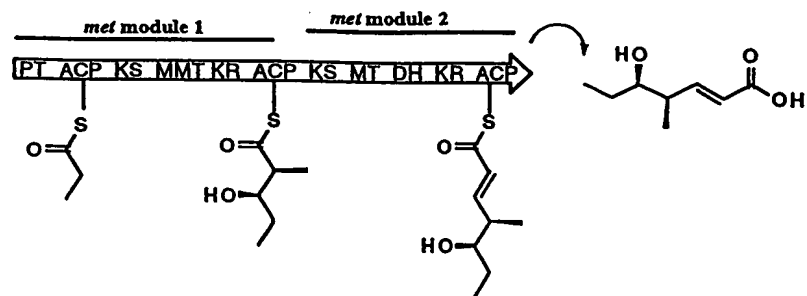


FIG. 5



1. introduce TE domain and establish release of acyl CoA ester
2. change MMT to MT domain in module 1
3. introduce DH/ER (or DH only) domain into module 1
4. inactivate DH domain in module 2
5. replace PT starter domain with AT in module 1

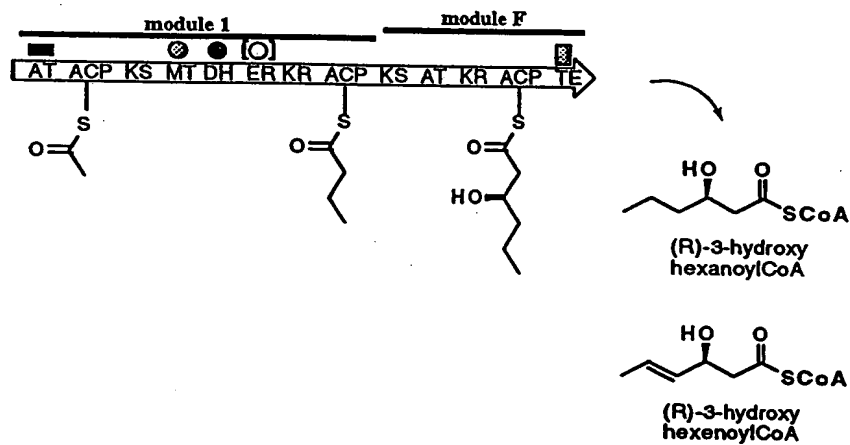


FIG. 6

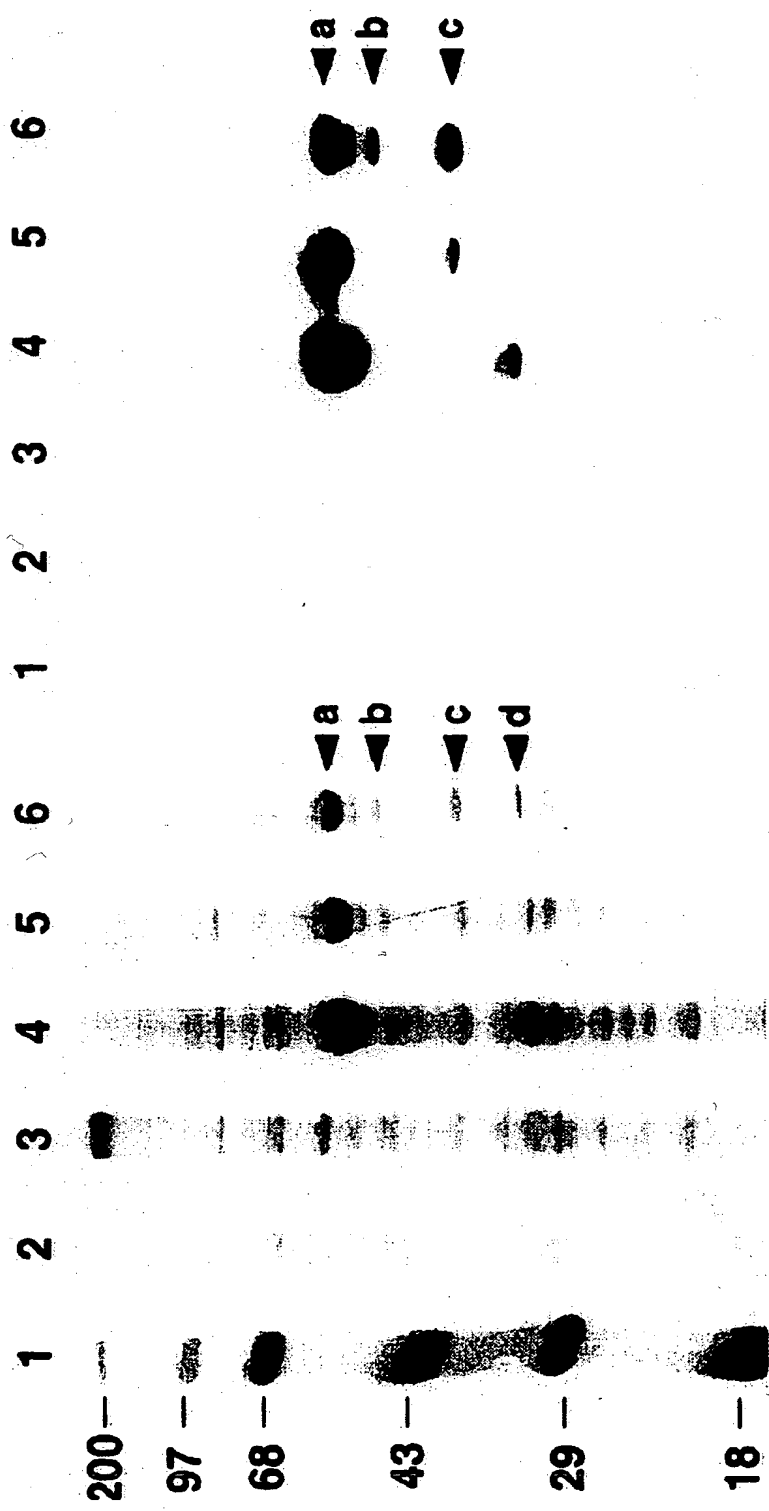


FIG. 7B

FIG. 7B

N-terminal sequence determined for PHA synthase

	1	10	20	25
a	MATGKGAAASTQEGKSQPFKVTPGP—			
b	AAASTQEGKSQPFKVTPGP—			
c	STQEGKSQPFKVTPGP—			

FIG. 8

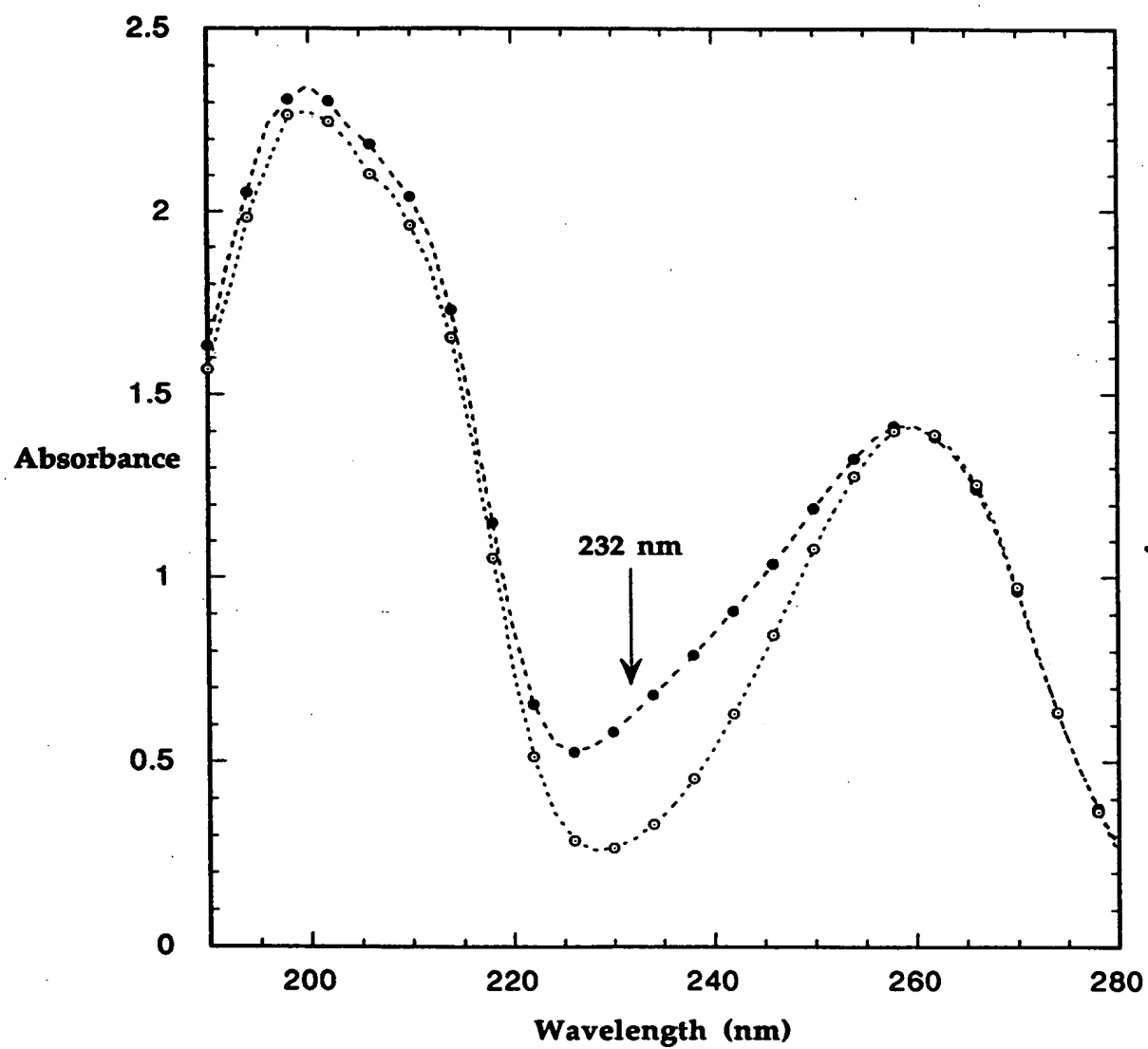


FIG. 9

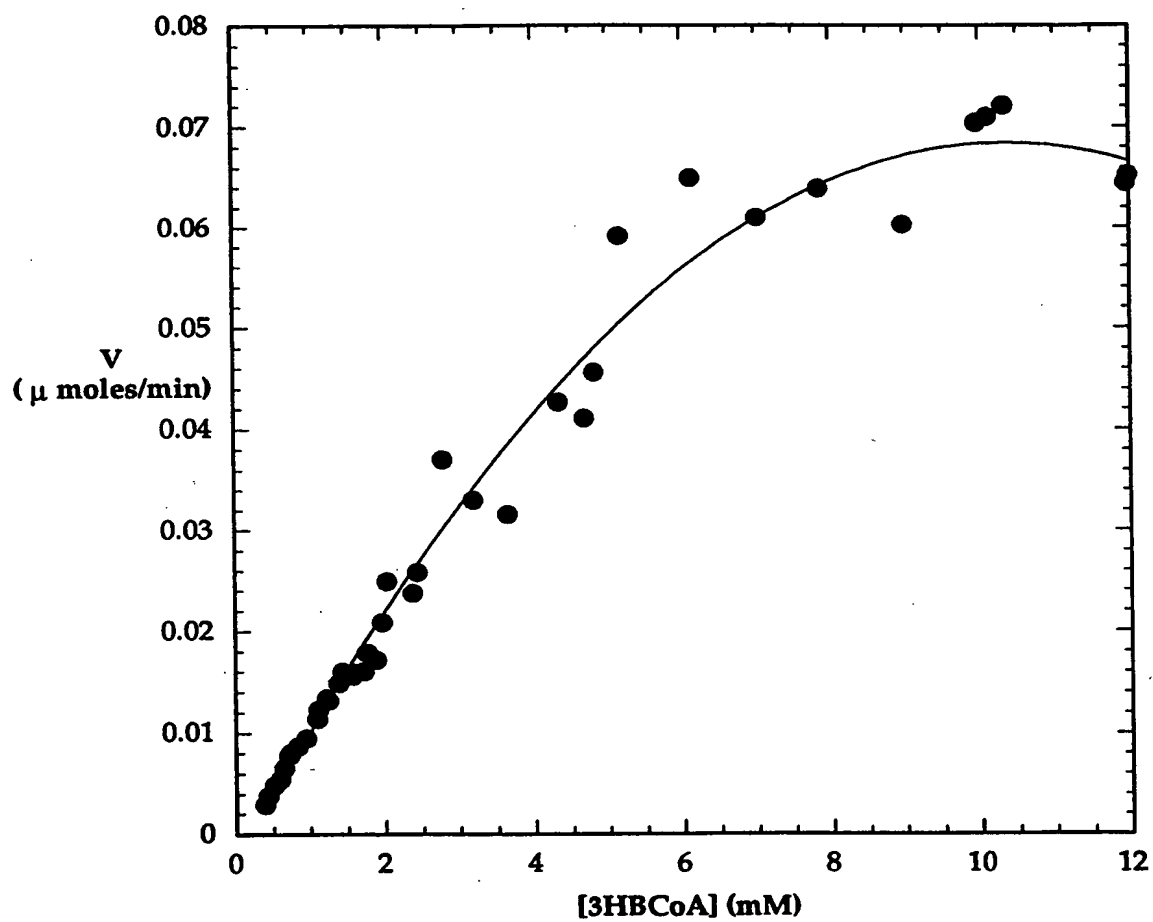


FIG. 10

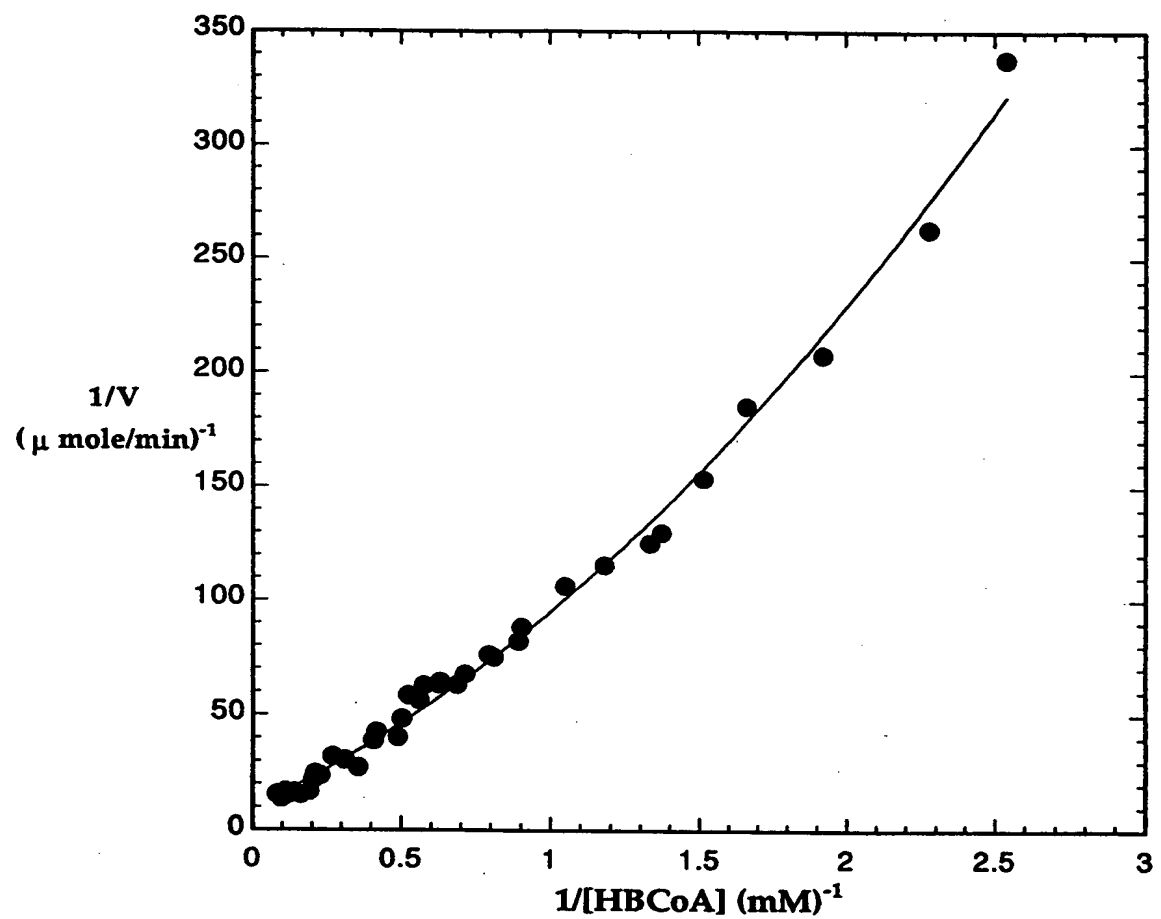


FIG. 11

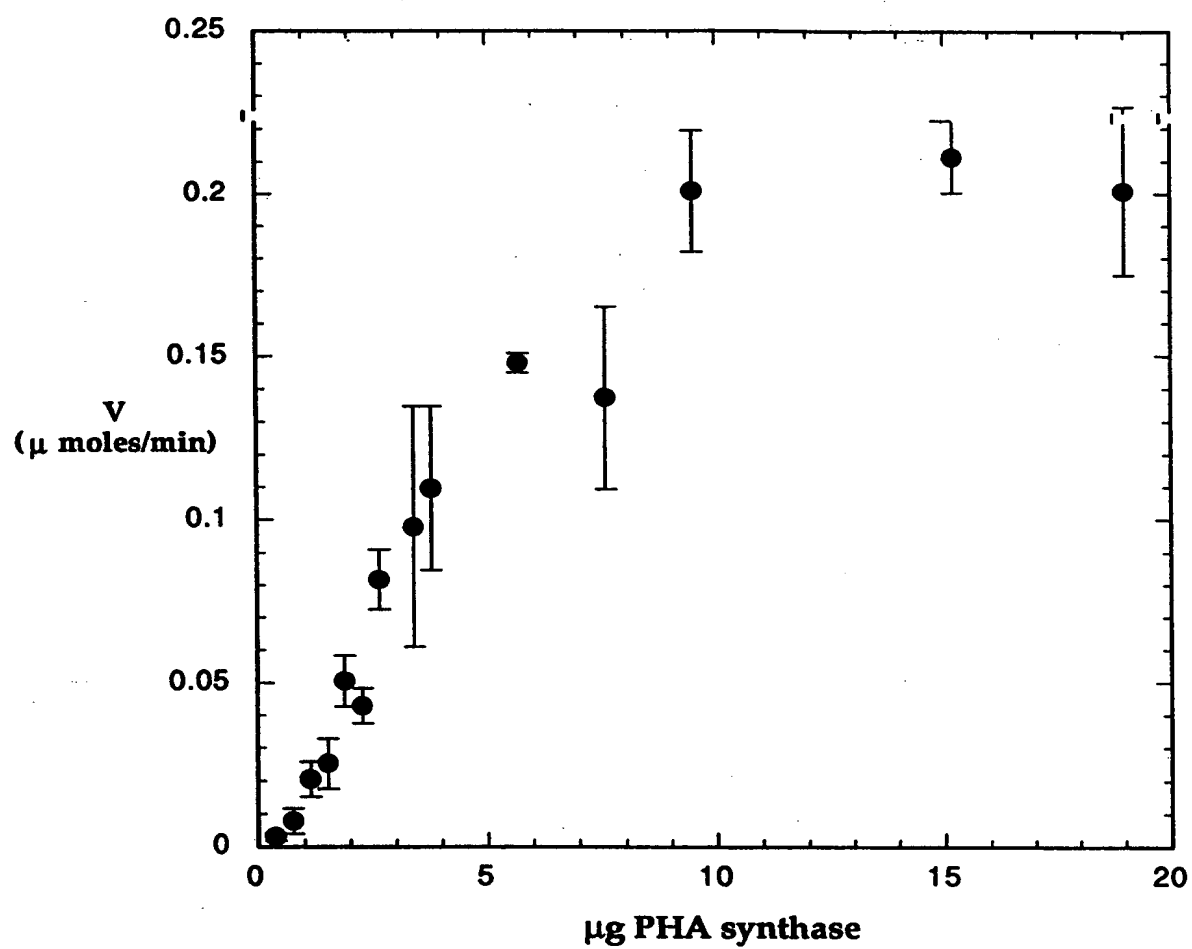


FIG. 12

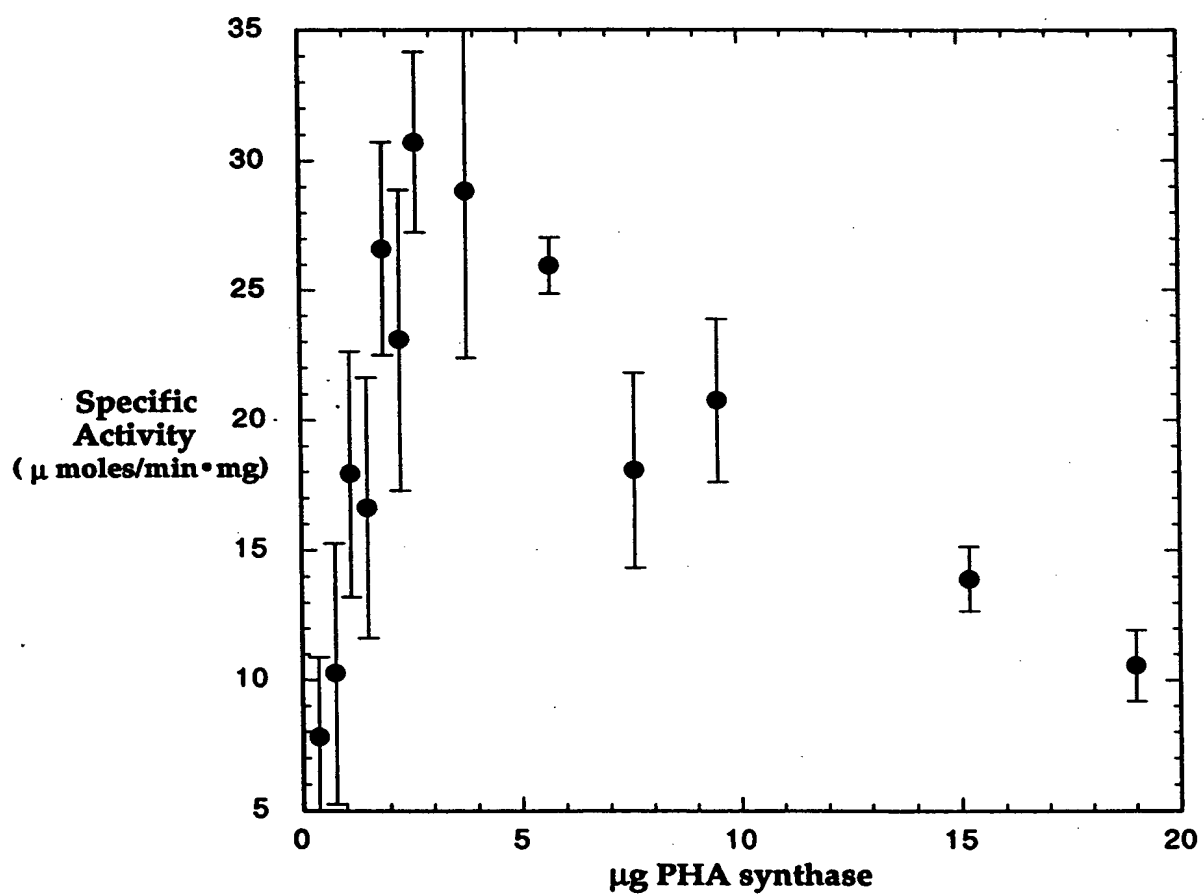


FIG. 13

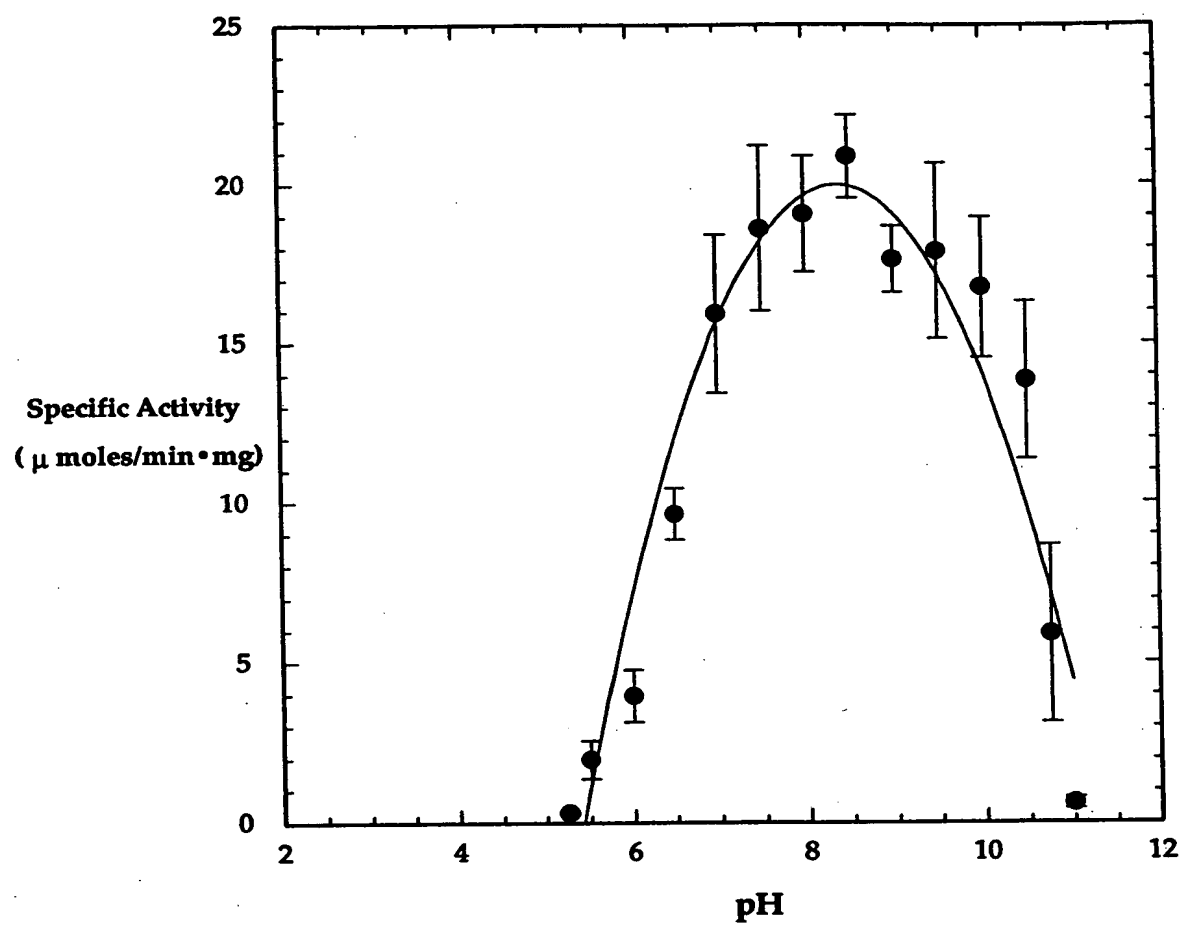


FIG. 14

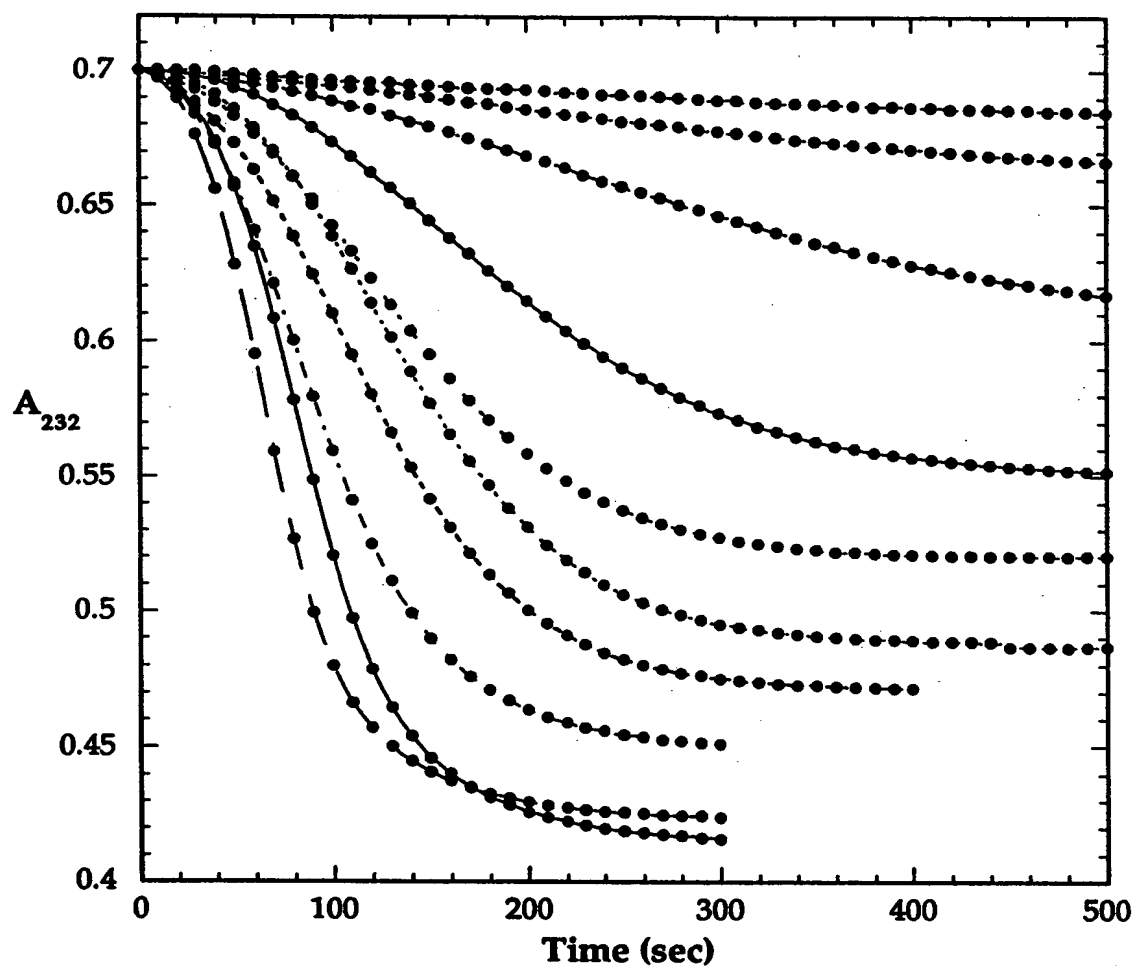


FIG. 15

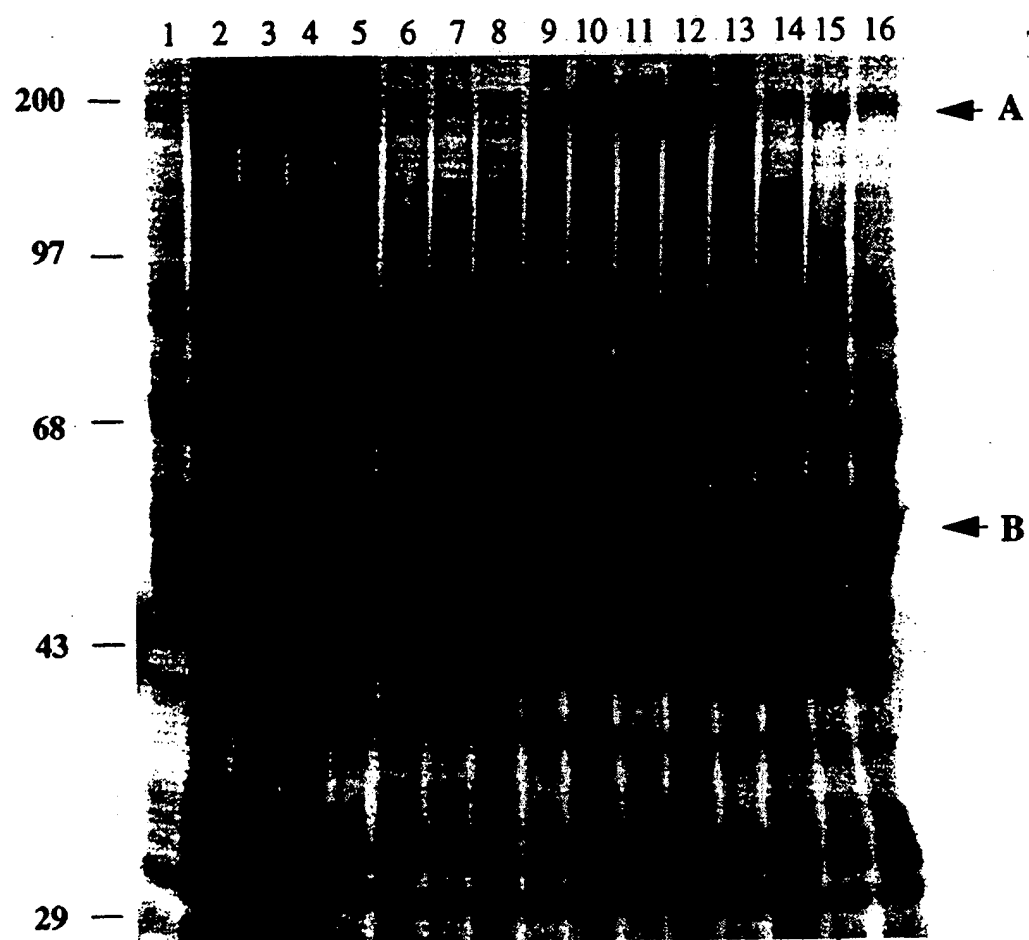


FIG. 16

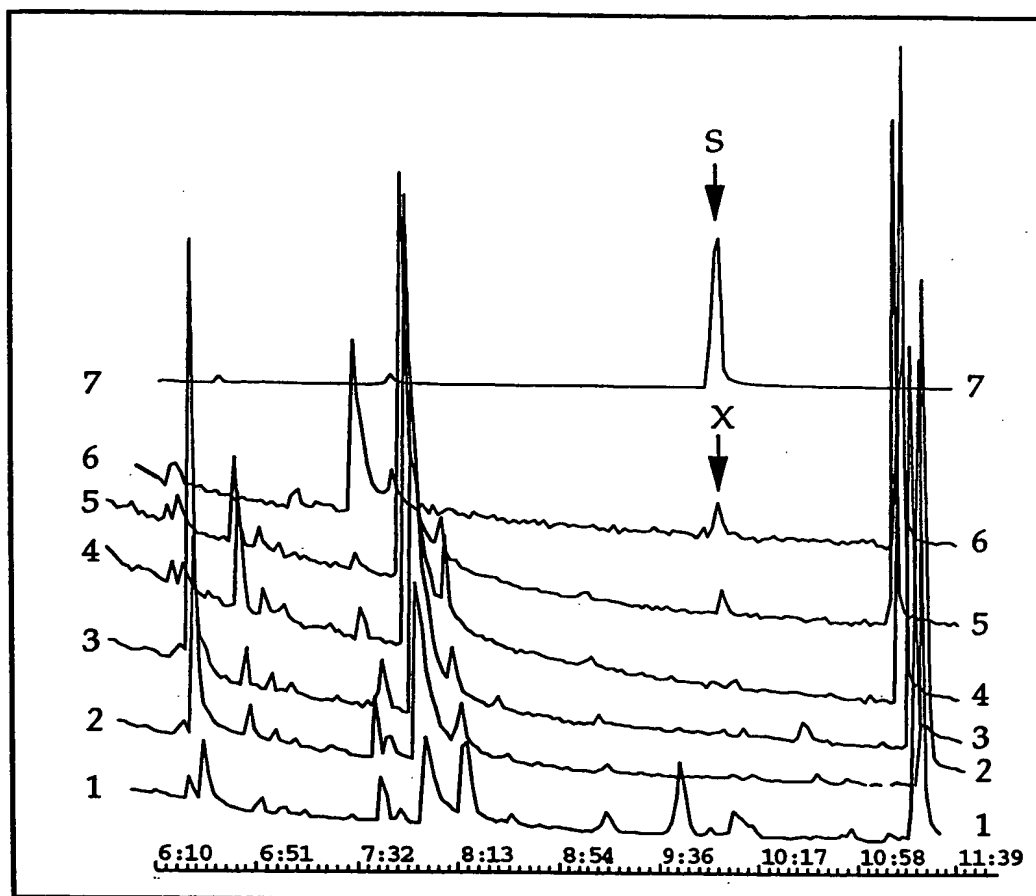


FIG. 17

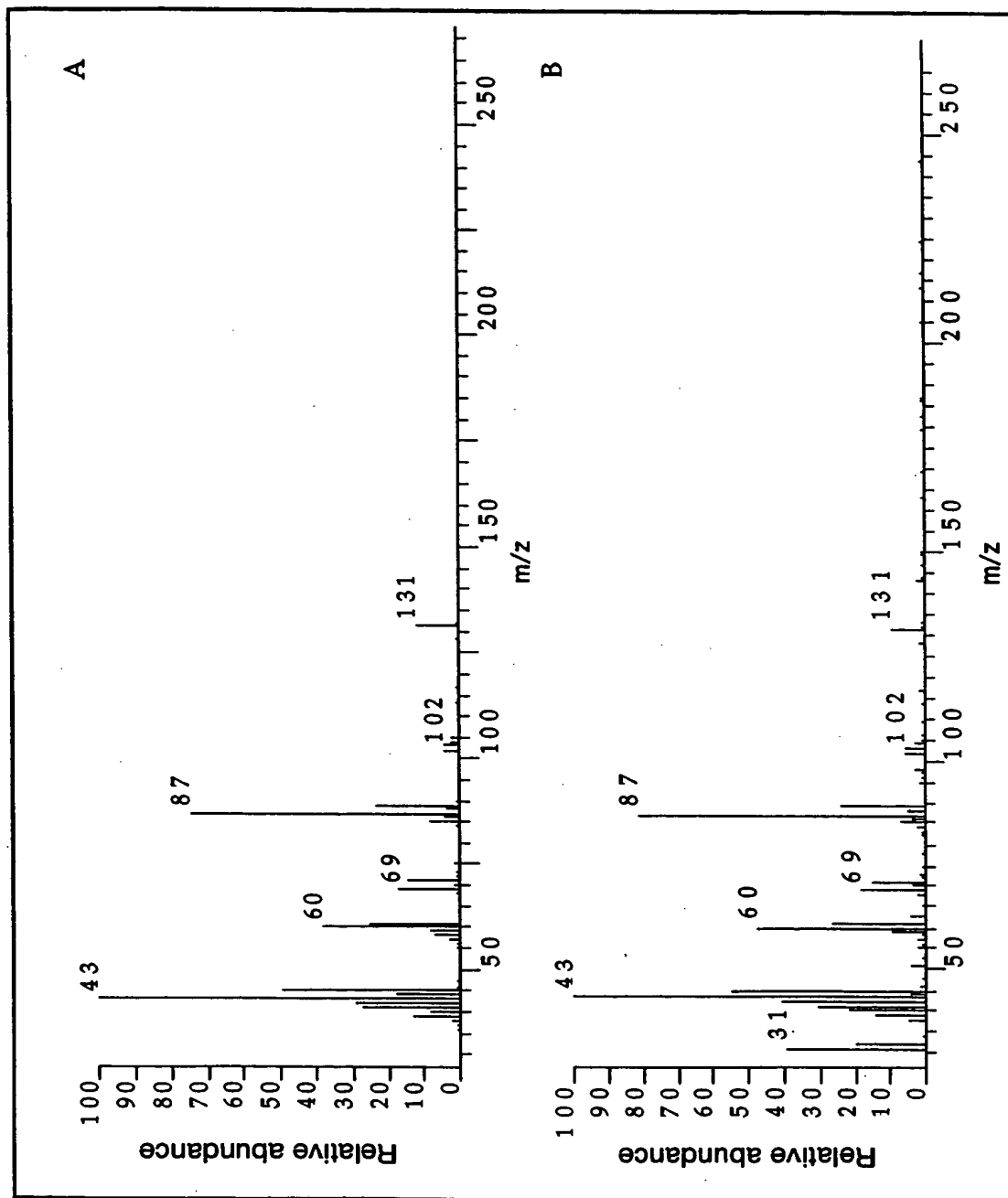


FIG. 18

[illegible]

Loading

==

 \geq

>

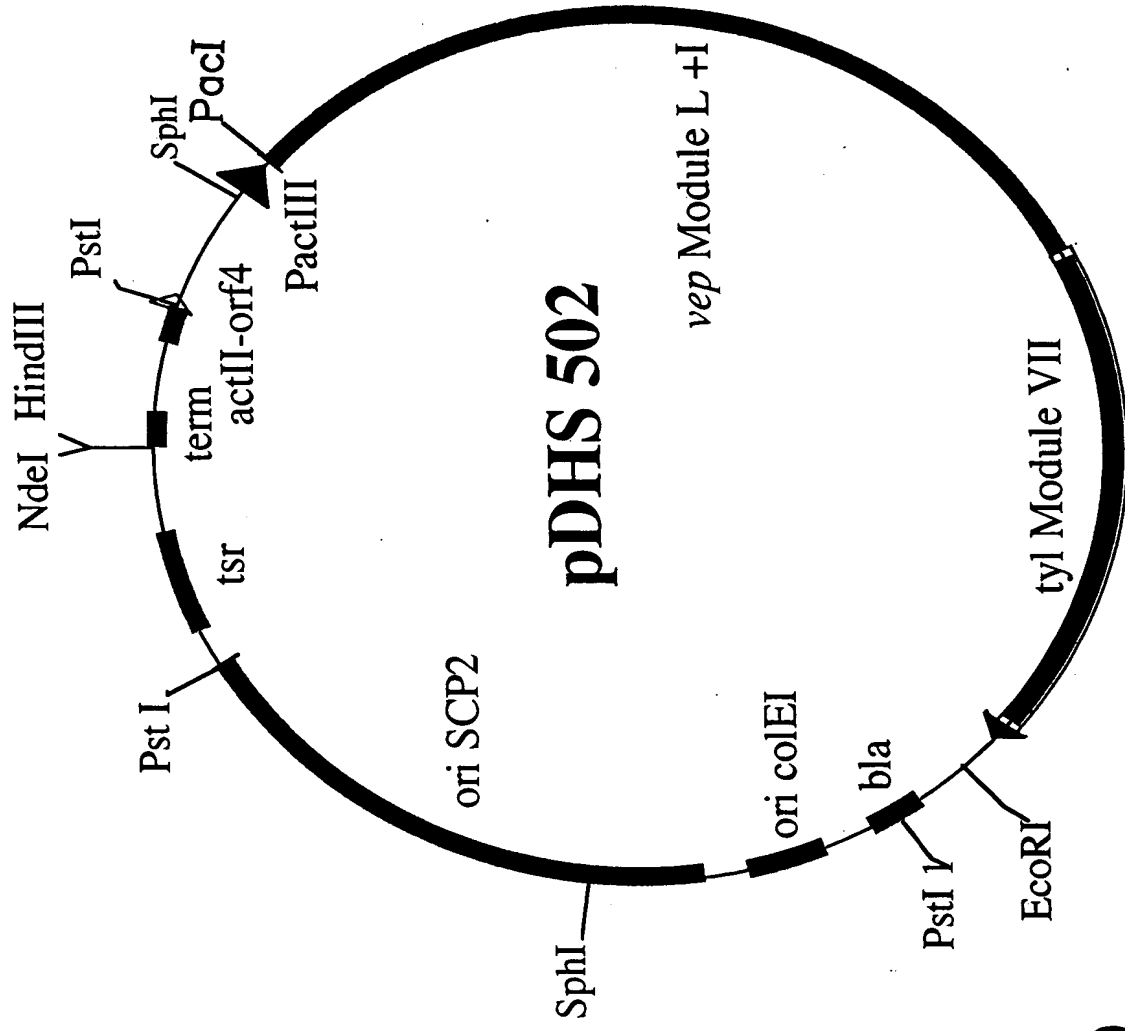
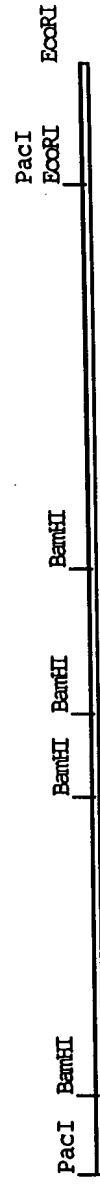


FIG. 20

pDHS505 Construction Procedure

Restriction map of pDHS505 insert

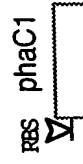
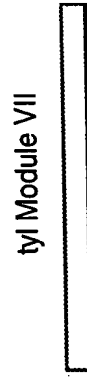


Major steps in the construction

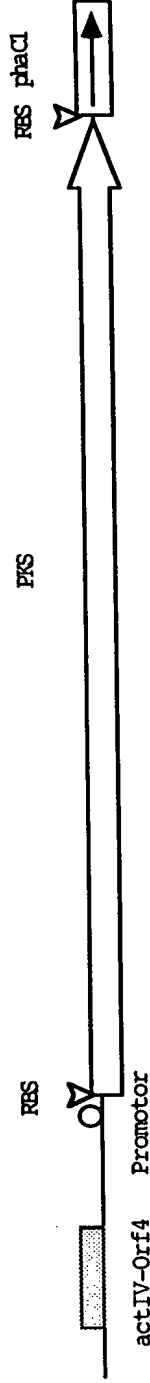
1. A *Streptomyces* Ribosome Binding Site (RBS) was introduced 6 nucleotides upstream of the translation start site to enhance gene translation in the host.



2. The *tyl* Module VII was recombined by a BamHI site with the *vcp* Module I to give out a complete polyketide synthase Open Reading Frame (ORF) with a Thioesterase at 3'-end.



3. The *phaC1* gene was transcriptionally coupled with the *PKS* gene. The second Ribosome Binding Site (RBS) was introduced to facilitate the gene translation.



4. The whole expression construct was put under the control of *act* promoter and the *actIV-Orf4* provides an activator which enhances the transcription and expression of the genes.

1	TTAATTAAGGAGGACCATC	ATG AAC GAG GCC ATC GCC GTC GTC GGC ATG TCC TGC CGC CTG CCG	64
1		M N E A I A V V G M S C R L P	15
65	AAG GCC TCG AAC CCG GCC GCC TTC TGG GAG CTG CTG CCG AAC GGG GAG AGC GCC GTC ACC	124	
16	K A S N P A A F W E L L R N G E S A V T	35	
125	GAC GTG CCC TCC GGC CCG TGG ACG TCG GTG CTC GGG GGA GCG GAC GCC GAG GAG CCG GCG	184	
36	D V P S G R W T S V L G G A D A E E P A	55	
185	GAG TCC GGT GTC CGC CCG GGC GGC TTC CTC GAC TCC CTC GAC CTC TTC GAC GCG GCC TTC	244	
56	E S G V R R G G F L D S L D L F D A A F	75	
245	TTC GGA ATC TCG CCC CGT GAG GCC GCC GCC ATG GAC CCG CAG CAG CGA CTG GTC CTC GAA	304	
76	F G I S P R E A A A M D P Q Q R L V L E	95	
305	CTC GCC TGG GAG GCG CTG GAG GAC GCC GGA ATC GTC CCC GGC ACC CTC GCC GGA AGC CCG	364	
96	L A W E A L E D A G I V P G T L A G S R	115	
365	ACC GCC GTC TTC GTC GGC ACC CTG CCG GAC GAC TAC ACG AGC CTC CTC TAC CAG CAC GGC	424	
116	T A V F V G T L R D D Y T S L L Y Q H G	135	
425	GAG CAG GCC ATC ACC CAG CAC ACC ATG GCG GGC GTG AAC CCG GGC GTC ATC GCC AAC CCG	484	
136	E Q A I T Q H T M A G V N R G V I A N R	155	
485	GTC TCG TAC CAC CTC GGC CTG CAG GGC CCG AGC CTC ACC GTC GAC GCC GCG CAG TCG TCC	544	
156	V S Y H L G L Q G P S L T V D A A Q S S	175	
545	TCG CTC GTC GCC GTG CAC CTG GCC TGC GAG TCC CTG CCG GCC GGG GAG TCC ACG ACG GCG	604	
176	S L V A V H L A C E S L R A G E S T T A	195	
605	CTC GTC GCC GGC GTG AAC CTC AAC ATC CTC GCG GAG AGC GCC GTG ACG GAG GAG CCG TTC	664	
196	L V A G V N L N I L A E S A V T E E R F	215	
665	GGT GGA CTC TCC CCG GAC GGC ACC GCC TAC ACC TTC GAC GCG CCG GCC AAC GGA TTC GTC	724	
216	G G L S P D G T A Y T F D A R A N G F V	235	
725	CGG GGC GAG GGC GGC GGA GTC GTC GTA CTC AAG CCG CTC TCC CCG GCC CTC GCC GAC GGC	784	
236	R G E G G G V V V L K P L S R A L A D G	255	
785	GAC CGT GTC CAC GGC GTC ATC CGC GCC AGC GCC GTC AAC AAC GAC GGA GCC ACC CCG GGT	844	
256	D R V H G V I R A S A V N N D G A T P G	275	
845	CTC ACC GTG CCC AGC AGG GCC GCC CAG GAG AAG GTG CTG CCG GAG GCG TAC CCG AAG GCG	904	
276	L T V P S R A A Q E K V L R E A Y R K A	295	
905	GCC CTG GAC CCG TCC GCC GTC CAG TAC GTC GAA CTC CAC GGC ACC GGA ACC CCC GTC GGC	964	
296	A L D P S A V Q Y V E L H G T G T P V G	315	
965	GAC CCC ATC GAG GCC GCC GCG CTC GGC GCC GTC CTC GGC TCG GCG CCG CCC GCG GAC GAA	1024	
316	D P I E A A A L G A V L G S A R P A D E	335	
1025	CCC CTG CTC GTC GGC TCG GCC AAG ACG AAC GTC GGG CAC CTC GAA GGC GCC GCC GGC ATC	1084	
336	P L L V G S A K T N V G H L E G A A G I	355	
1085	GTC GGC CTC ATC AAG ACG CTC CTC GCG CTC GGC CCG CCG CCG ATC CCG GCG AGC CTC AAC	1144	
356	V G L I K T L L A L G R R R I P A S L N	375	
1145	TTC CGT ACG CCC CAC CCG GAC ATC CCG CTC GAC ACC CTC GGG CTC GAC GTG CCC GAC GGC	1204	
376	F R T P H P D I P L D T L G L D V P D G	395	
1205	CTG CCG GAG TGG CCG CAC CCG GAC CCG GAA CTC CTC GCC GGC GTC AGC TCG TTC GGC ATG	1264	
396	L R E W P H P D R E L L A G V S S F G M	415	
1265	GGC GGC ACC AAC GCC CAC GTC GTC CTC AGC GAA GGC CCC GCC CAG GGC GGC GAG CAG CCC	1324	
416	G G T N A H V V L S E G P A Q G G E Q P	435	
1325	GGC ATC GAT GAG GAG ACC CCC GTC GAC AGC GGG GCC GCA CTG CCC TTC GTC GTC ACC GGC	1384	
436	G I D E E T P V D S G A A L P F V V T G	455	
1385	CGC GGC GGC GAG GCC CTG CCG GCC CAG GCC CCG CCG CTG CAC GAG GCC GTC GAA GCG GAC	1444	
456	R G G E A L R A Q A R R L H E A V E A D	475	

FIG. 23A

1445	CCG	GAG	CTC	GCG	CCC	GCC	GCA	CTC	GCC	CGG	TCG	CTG	GTC	ACC	ACC	CGT	ACG	GTC	TTC	ACG	1504
476	P	E	L	A	P	A	A	L	A	R	S	L	V	T	T	R	T	V	F	T	495
1505	CAC	CGG	TCG	GTC	GTC	CTC	GCC	CCG	GAC	CGC	GCC	CGC	CTC	CTC	GAC	GGC	CTC	GGC	GCC	CTC	1564
496	H	R	S	V	V	L	A	P	D	R	A	R	L	L	D	G	L	G	A	L	515
1565	GCC	GCC	GGG	ACG	CCC	GCG	CCC	GGC	GTG	GTC	ACC	GGC	ACC	CCC	GCC	CCC	GGG	CGC	CTC	GCC	1624
516	A	A	G	T	P	A	P	G	V	V	T	G	T	P	A	P	G	R	L	A	535
1625	GTC	CTG	TTC	AGC	GGC	CAG	GGT	GCC	CAA	CGT	ACG	GGC	ATG	GGC	ATG	GAG	TTG	TAC	GCC	GCC	1684
536	V	L	F	S	G	Q	G	A	Q	R	T	G	M	G	M	E	L	Y	A	A	555
1685	CAC	CCC	GCC	TTC	GCG	ACG	GCC	TTC	GAC	GCC	GTC	GCC	GCC	GAA	CTG	GAC	CCC	CTC	CTC	GAC	1744
556	H	P	A	F	A	T	A	F	D	A	V	A	A	E	L	D	P	L	L	D	575
1745	CGG	CCC	CTC	GCC	GAA	CTC	GTC	GCG	GCG	GGC	GAC	ACC	CTC	GAC	CGC	ACC	GTC	CAC	ACA	CAG	1804
576	R	P	L	A	E	L	V	A	A	G	D	T	L	D	R	T	V	H	T	Q	595
1805	CCC	GCG	CTC	TTC	GCC	GTG	GAG	GTC	GCC	CTC	CAC	CGC	CTC	GTC	GAG	TCC	TGG	GGC	GTC	ACG	1864
596	P	A	L	F	A	V	E	V	A	L	H	R	L	V	E	S	W	G	V	T	615
1865	CCC	GAC	CTG	CTC	GCC	GGC	CAC	TCC	GTC	GGC	GAG	ATC	AGC	GCC	GCC	CAC	GTC	GCC	GGG	GTC	1924
616	P	D	L	L	A	G	H	S	V	G	E	I	S	A	A	H	V	A	G	V	635
1925	CTG	TCG	CTG	CGC	GAC	GCC	GCC	CGC	CTC	GTC	GCG	GCG	CGC	GGC	CGC	CTC	ATG	CAG	GCG	CTC	1984
636	L	S	L	R	D	A	A	R	L	V	A	A	R	G	R	L	M	Q	A	L	655
1985	CCC	GAG	GGC	GGC	GCG	ATG	GTC	GCG	GTC	GAG	GCG	AGC	GAG	GAG	GAA	GTG	CTT	CCG	CAC	CTC	2044
656	P	E	G	G	A	M	V	A	V	E	A	S	E	E	E	V	L	P	H	L	675
2045	GCG	GGA	CGC	GAG	CGG	GAG	CTC	TCC	CTC	GCG	GCC	GTG	AAC	GGC	CCC	CGC	GCG	GTC	GTC	CTC	2104
676	A	G	R	E	R	E	L	S	L	A	A	V	N	G	P	R	A	V	V	L	695
2105	GCG	GGC	GCC	GAG	CGC	GCC	GTC	CTC	GAC	GTC	GCC	GAG	CTG	CTG	CGC	GAA	CAG	GGC	CGC	CGG	2164
696	A	G	A	E	R	A	V	L	D	V	A	E	L	L	R	E	Q	G	R	R	715
2165	ACG	AAG	CGG	CTC	AGC	GTC	TCG	CAC	GCC	TTC	CAC	TCG	CCG	CTC	ATG	GAG	CCG	ATG	CTC	GAC	2224
716	T	K	R	L	S	V	S	H	A	F	H	S	P	L	M	E	P	M	L	D	735
2225	GAC	TTC	CGC	CGG	GTC	GTC	GAA	GAG	CTG	GAC	TTC	CAG	GAG	CCC	CGC	GTC	GAC	GTC	GTG	TCC	2284
736	D	F	R	R	V	V	E	E	L	D	F	Q	E	P	R	V	D	V	V	S	755
2285	ACG	GTG	ACG	GGC	CTG	CCT	GTC	ACA	GCG	GGC	CAA	TGG	ACC	GAT	CCC	GAG	TAC	TGG	GTG	GAC	2344
756	T	V	T	G	L	P	V	T	A	G	Q	W	T	D	P	E	Y	W	V	D	775
2345	CAG	GTC	CGC	AGG	CCC	GTA	CGC	TTC	CTC	GAC	GCC	GTA	CGC	ACC	CTG	GAG	GAA	TCG	GGC	GCC	2404
776	Q	V	R	R	P	V	R	F	L	D	A	V	R	T	L	E	E	S	G	A	795
2405	GAC	ACC	TTC	CTG	GAG	CTC	GGT	CCC	GAC	GGG	GTC	TGC	TCC	GCG	ATG	GCG	GCG	GAC	TCC	GTA	2464
796	D	T	F	L	E	L	G	P	D	G	V	C	S	A	M	A	A	D	S	V	815
2465	CGC	GAC	CAG	GAG	GCC	GCC	ACG	GCG	GTC	TCC	GCC	CTG	CGC	AAG	GGC	CGC	CCG	GAG	CCC	CAG	2524
816	R	D	Q	E	A	A	T	A	V	S	A	L	R	K	G	R	P	E	P	Q	835
2525	TCG	CTG	CTC	GCC	GCA	CTC	ACC	ACC	GTC	TTC	GTC	CGG	GGC	CAC	GAC	GTC	GAC	TGG	ACC	GCC	2584
836	S	L	L	A	A	L	T	T	V	F	V	R	G	H	D	V	D	W	T	A	855
2585	GCG	CAC	GGG	AGC	ACC	GGC	ACG	GTC	AGG	GTG	CCC	CTG	CCG	ACC	TAC	GCC	TTC	CAG	CGC	GAA	2644
856	A	H	G	S	T	G	T	V	R	V	P	L	P	T	Y	A	F	Q	R	E	875
2645	CGC	CAC	TGG	TTC	GAC	GGC	GCC	GCG	CGA	ACG	GCG	GCC	CCG	CTC	ACG	GCG	GGC	CGA	TCG	GGC	2704
876	R	H	W	F	D	G	A	A	R	T	A	A	P	L	T	A	G	R	S	G	895
2705	ACC	GGT	GCG	GGC	ACC	GGC	CCG	GCC	GCG	GGT	GTG	ACG	TCG	GGC	GAG	GGC	GAG	GGC	GAG	GGC	2764
896	T	G	A	G	T	G	P	A	A	G	V	T	S	G	E	G	E	G	E	G	915
2765	GAG	GGC	GCG	GGT	GCG	GGT	GGC	GGT	GAT	CGG	CCG	GCT	CGC	CAC	GAG	ACG	ACC	GAG	CGC	GTG	2824
916	E	G	A	G	A	G	G	G	D	R	P	A	R	H	E	T	T	E	R	V	935
2825	CGC	GCA	CAC	GTC	GCC	GCC	GTC	CTC	GAG	TAC	GAC	GAC	CCG	ACC	CGC	GTC	GAA	CTC	GGC	CTC	2884
936	R	A	H	V	A	A	V	L	E	Y	D	D	P	T	R	V	E	L	G	L	955
2885	ACC	TTC	AAG	GAG	CTG	GGC	TTC	GAC	TCC	CTC	ATG	TCC	GTC	GAG	CTG	CGG	AAC	GCG	CTC	GTC	2944
956	T	F	K	E	L	G	F	D	S	L	M	S	V	E	L	R	N	A	L	V	975
2945	GAC	GAC	ACG	GGA	CTG	CGC	CTG	CCC	AGC	GGA	CTG	CTC	TTC	GAC	CAC	CCG	ACG	CCG	CGC	GCC	3004
976	D	D	T	G	L	R	L	P	S	G	L	L	F	D	H	P	T	P	R	A	995

FIG. 23B

3005	CTC	GCC	GCC	CAC	CTG	GGC	GAC	CTG	CTC	ACC	GGC	GGC	AGC	GGC	GAG	ACC	GGA	TCG	GCC	GAC	3064
996	L	A	A	H	L	G	D	L	L	T	G	G	S	G	E	T	G	S	A	D	1015
3065	GGG	ATA	COG	CCC	GCG	ACC	COG	GCG	GAC	ACC	ACC	GCC	GAG	CCC	ATC	GCG	ATC	ATC	GGC	ATG	3124
1016	G	I	P	P	A	T	P	A	D	T	T	A	E	P	I	A	I	I	G	M	1035
3125	GCC	TGC	CGC	TAC	CCC	GGC	GGC	GTC	ACC	TCC	CCC	GAG	GAC	CTG	TGG	CGG	CTC	GTC	GCC	GAG	3184
1036	A	C	R	Y	P	G	G	V	T	S	P	E	D	L	W	R	L	V	A	E	1055
3185	GGG	CGC	GAC	GCC	GTC	TCG	GGG	CTG	CCC	ACC	GAC	CGC	GGC	TGG	GAC	GAG	GAC	CTC	TTC	GAC	3244
1056	G	R	D	A	V	S	G	L	P	T	D	R	G	W	D	E	D	L	F	D	1075
3245	GCC	GAC	CCC	GAC	CGC	AGC	GGC	AAG	AGC	TCG	GTC	CGC	GAG	GGC	GGA	TTC	CTG	CAC	GAC	GCC	3304
1076	A	D	P	D	R	S	G	K	S	S	V	R	E	G	G	F	L	H	D	A	1095
3305	GCC	CTG	TTC	GAC	GCC	GGC	TTC	TTC	GGG	ATA	TCG	CCC	CGC	GAG	GCC	CTC	GGC	ATG	GAC	CCG	3364
1096	A	L	F	D	A	G	F	F	G	I	S	P	R	E	A	L	G	M	D	P	1115
3365	CAG	CAG	CGG	CTG	CTC	CTG	GAG	ACG	GCA	TGG	GAG	GCC	GTC	GAG	CGC	GCA	GGG	CTC	GAC	CCC	3424
1116	Q	Q	R	L	L	L	E	T	A	W	E	A	V	E	R	A	G	L	D	P	1135
3425	GAA	GGC	CTC	AAG	GGC	AGC	CGG	ACG	GCC	GTC	TTC	GTC	GGC	GCC	ACC	GCC	CTG	GAC	TAC	GGC	3484
1136	E	G	L	K	G	S	R	T	A	V	F	V	G	A	T	A	L	D	Y	G	1155
3485	COG	CGC	ATG	CAC	GAC	GGC	GCC	GAG	GGC	GTC	GAG	GGC	CAC	CTC	CTG	ACC	GGG	ACC	ACG	CCC	3544
1156	P	R	M	H	D	G	A	E	G	V	E	G	H	L	L	T	G	T	T	P	1175
3545	AGC	GTG	ATG	TCG	GGC	CGC	ATC	GCC	TAC	CAG	CTC	GGC	CTC	ACC	GGT	CCT	GCG	GTC	ACC	GTC	3604
1176	S	V	M	S	G	R	I	A	Y	Q	L	G	L	T	G	P	A	V	T	V	1195
3605	GAC	ACG	GCC	TGC	TCG	TCC	TCG	CTC	GTC	GCG	CTG	CAC	CTG	GCC	GTC	CGT	TCG	CTG	CGG	CAG	3664
1196	D	T	A	C	S	S	S	L	V	A	L	H	L	A	V	R	S	L	R	Q	1215
3665	GGC	GAG	TCG	AGC	CTC	GCG	CTC	GCC	GGC	GGA	GCG	ACC	GTC	ATG	TCG	ACA	CCG	GGC	ATG	TTC	3724
1216	G	E	S	S	L	A	L	A	G	G	A	T	V	M	S	T	P	G	M	F	1235
3725	GTC	GAG	TTC	TCG	CGG	CAG	CGC	GGC	CTC	GCC	GCC	GAC	GGC	CGC	TCC	AAG	GCC	TTC	TCC	GAC	3784
1236	V	E	F	S	R	Q	R	G	L	A	A	D	G	R	S	K	A	F	S	D	1255
3785	TCC	GCC	GAC	GGC	ACC	TCC	TGG	GCC	GAG	GGC	GTC	GGC	CTC	CTC	GTC	GTC	GAG	CGG	CTC	TCG	3844
1256	S	A	D	G	T	S	W	A	E	G	V	G	L	L	V	V	E	R	L	S	1275
3845	GAC	GCC	GAG	CGC	AAC	GGC	CAC	CCC	GTG	CTC	GCC	GTG	ATC	CGG	GGC	AGC	GCG	GTC	AAC	CAG	3904
1276	D	A	E	R	N	G	H	P	V	L	A	V	I	R	G	S	A	V	N	Q	1295
3905	GAC	GGC	GCC	TCC	AAC	GGG	CTC	ACC	GCC	CCC	AAC	GGC	COG	TCC	CAG	CAG	CGC	GTC	ATC	CGA	3964
1296	D	G	A	S	N	G	L	T	A	P	N	G	P	S	Q	Q	R	V	I	R	1315
3965	CAG	GCC	CTG	GCC	GAC	GCC	GGG	CTC	ACC	COG	GCC	GAC	GTC	GAC	GCC	GTC	GAG	GCG	CAC	GGT	4024
1316	Q	A	L	A	D	A	G	L	T	P	A	D	V	D	A	V	E	A	H	G	1335
4025	ACG	GGT	ACC	CGG	CTC	GGC	GAC	CCC	ATC	GAG	GCC	GAG	GCG	ATC	CTC	GGC	ACC	TAC	GGC	CGG	4084
1336	T	G	T	R	L	G	D	P	I	E	A	E	A	I	L	G	T	Y	G	R	1355
4085	GAC	CGG	GGC	GAG	GGC	GCT	COG	CTC	CAG	CTC	GGC	TCG	CTG	AAG	TCG	AAC	ATC	GGC	CAC	GCG	4144
1356	D	R	G	E	G	A	P	L	Q	L	G	S	L	K	S	N	I	G	H	A	1375
4145	CAG	GCC	GCC	GCG	GGC	GTG	GGC	GGG	CTC	ATC	AAG	ATG	GTC	CTC	GCG	ATG	CGC	CAC	GGC	GTC	4204
1376	Q	A	A	A	G	V	G	G	L	I	K	M	V	L	A	M	R	H	G	V	1395
4205	CTG	CCC	AGG	ACG	CTC	CAC	GTG	GAC	CGG	CCC	ACC	ACC	CGC	GTC	GAC	TGG	GAG	GCC	GGC	GGC	4264
1396	L	P	R	T	L	H	V	D	R	P	T	T	R	V	D	W	E	A	G	G	1415
4265	GTC	GAG	CTC	CTC	ACC	GAG	GAG	CGG	GAG	TGG	CCG	GAG	ACG	GGC	CGC	CCG	CGC	CGC	GCG	GCG	4324
1416	V	E	L	L	T	E	E	R	E	W	P	E	T	G	R	P	R	R	A	A	1435
4325	ATC	TCC	TCC	TTC	GGC	ATC	AGC	GGC	ACC	AAC	GCC	CAC	ATC	GTG	GTC	GAA	CAG	GCC	CCG	GAA	4384
1436	I	S	S	F	G	I	S	G	T	N	A	H	I	V	V	E	Q	A	P	E	1455
4385	GCC	GGG	GAG	GCG	GCG	GTC	ACC	ACC	ACC	GCC	CCG	GAA	GCA	GGG	GAA	GCC	GGG	GAA	GCG	GCG	4444
1456	A	G	E	A	A	V	T	T	T	A	P	E	A	G	E	A	G	E	A	A	1475
4445	GAC	ACC	ACC	GCC	ACC	ACG	ACG	COG	GCC	GCG	GTC	GGC	GTC	CCC	GAA	CCC	GTA	CGC	GCC	CCC	4504
1476	D	T	T	A	T	T	T	P	A	A	V	G	V	P	E	P	V	R	A	P	1495
4505	GTC	GTG	GTC	TCC	GCG	CGG	GAC	GCC	GCC	GCC	CTG	CGC	GCC	CAG	GCC	GTT	CGG	CTG	CGG	ACC	4564
1496	V	V	V	S	A	R	D	A	A	A	L	R	A	Q	A	V	R	L	R	T	1515

FIG. 23C

4565	TTC	CTC	GAC	GGC	CGA	CCG	GAC	GTC	ACC	GTC	GCC	GAC	CTC	GGA	CGC	TCG	CTG	GCC	GCC	CGT	4624
1516	F	L	D	G	R	P	D	V	T	V	A	D	L	G	R	S	L	A	A	R	1535
4625	ACC	GCC	TTC	GAG	CAC	AAG	GCC	GCC	CTC	ACC	ACC	GCC	ACC	AGG	GAC	GAG	CTG	CTC	GCC	GGG	4684
1536	T	A	F	E	H	K	A	A	L	T	T	A	T	R	D	E	L	L	A	G	1555
4685	CTC	GAC	GCC	CTC	GGC	CGC	GGG	GAG	CAA	GCC	ACG	GGC	CTG	GTC	ACC	GGC	GAA	CCG	GCC	AGG	4744
1556	L	D	A	L	G	R	G	E	Q	A	T	G	L	V	T	G	E	P	A	R	1575
4745	GCC	GGA	CGC	ACG	GCC	TTC	CTG	TTC	ACC	GGC	CAG	GGA	GCG	CAG	CGC	GTC	GCC	ATG	GGC	GAG	4804
1576	A	G	R	T	A	F	L	F	T	G	Q	G	A	Q	R	V	A	M	G	E	1595
4805	GAA	CTG	CGC	GCC	GCG	CAC	CCC	GTG	TTC	GCC	GCC	GCC	CTC	GAC	ACC	GTG	TAC	GCG	GCC	CTC	4864
1596	E	L	R	A	A	H	P	V	F	A	A	A	L	D	T	V	Y	A	A	L	1615
4865	GAC	CGT	CAC	CTC	GAC	CGG	CCG	CTG	CGG	GAG	ATC	GTC	GCC	GCC	GGG	GAG	GAG	CTG	GAC	CTC	4924
1616	D	R	H	L	D	R	P	L	R	E	I	V	A	A	G	E	E	L	D	L	1635
4925	ACC	GCG	TAC	ACC	CAG	CCC	GCC	CTC	TTC	GCC	TTC	GAG	GTG	GCG	CTG	TTC	CGC	CTC	CTC	GAA	4984
1636	T	A	Y	T	Q	P	A	L	F	A	F	E	V	A	L	F	R	L	L	E	1655
4985	CAC	CAC	GGC	CTC	GTC	CCC	GAC	CTG	CTC	ACC	GGC	CAC	TCC	GTG	GGC	GAG	ATC	GCC	GCC	GCG	5044
1656	H	H	G	L	V	P	D	L	L	T	G	H	S	V	G	E	I	A	A	A	1675
5045	CAC	GTC	GCC	GGT	GTC	CTC	TCC	CTC	GAC	GAC	GCC	GCA	CGT	CTC	GTC	ACC	GCC	CGC	GGC	CGG	5104
1676	H	V	A	G	V	L	S	L	D	D	A	A	R	L	V	T	A	R	G	R	1695
5105	CTC	ATG	CAG	TCG	GCC	CGC	GAG	GGC	GGC	GCG	ATG	ATC	GCC	GTG	CAG	GCG	GGC	GAG	GCC	GAG	5164
1696	L	M	Q	S	A	R	E	G	G	A	M	I	A	V	Q	A	G	E	A	E	1715
5165	GTC	GTC	GAG	TCC	CTG	AAG	GGC	TAC	GAG	GGC	AGG	GTC	GCC	GTG	GCC	GCC	GTC	AAC	GGA	CCC	5224
1716	V	V	E	S	L	K	G	Y	E	G	R	V	A	V	A	A	V	N	G	P	1735
5225	ACC	GCC	GTG	GTC	GTC	TCC	GGC	GAC	GCG	GAC	GCC	GCC	GAG	GAG	ATC	CGC	GCC	GTA	TGG	GCG	5284
1736	T	A	V	V	V	S	G	D	A	D	A	A	E	E	I	R	A	V	W	A	1755
5285	GGA	CGC	GGC	CGG	CGC	ACC	CGC	AGG	CTG	CGC	GTC	AGC	CAC	GCC	TTC	CAC	TCC	CCG	CAC	ATG	5344
1756	G	R	G	R	R	T	R	R	L	R	V	S	H	A	F	H	S	P	H	M	1775
5345	GAC	GAC	GTC	CTC	GAC	GAG	TTC	CTC	CGG	GTC	GCC	GAG	GGC	CTG	ACC	TTC	GAG	GAG	CCG	CGG	5404
1776	D	D	V	L	D	E	F	L	R	V	A	E	G	L	T	F	E	E	P	R	1795
5405	ATC	CCC	GTC	GTC	TCC	ACG	GTC	ACC	GGC	GCG	CTC	GTC	ACG	TCC	GGC	GAG	CTC	ACC	TCG	CCC	5464
1796	I	P	V	V	S	T	V	T	G	A	L	V	T	S	G	E	L	T	S	P	1815
5465	GCG	TAC	TGG	GTC	GAC	CAG	ATC	CGG	CGG	CCC	GTG	CGC	TTC	CTG	GAC	GCC	GTC	CGC	ACC	CTG	5524
1816	A	Y	W	V	D	Q	I	R	R	P	V	R	F	L	D	A	V	R	T	L	1835
5525	GCC	GCC	CAG	GAC	GCG	ACC	GTC	CTC	GTC	GAG	ATC	GGC	CCC	GAC	GCC	GTC	CTC	ACG	GCA	CTC	5584
1836	A	A	Q	D	A	T	V	L	V	E	I	G	P	D	A	V	L	T	A	L	1855
5585	GCC	GAG	GAG	GCT	CTC	GCG	CCC	GGC	ACG	GAC	GCC	CGG	GAC	GCC	CGG	GAC	GTC	ACG	GTC	GTC	5644
1856	A	E	E	A	L	A	P	G	T	D	A	P	D	A	R	D	V	T	V	V	1875
5645	CCG	CTG	CTG	CGC	GCG	GGG	CGC	CCC	GAG	CCC	GAG	ACC	CTC	GCC	GCC	GGT	CTC	GCG	ACC	GCC	5704
1876	P	L	L	R	A	G	R	P	E	P	E	T	L	A	A	G	L	A	T	A	1895
5705	CAT	GTC	CAC	GGC	GCA	CCC	TTG	GAC	CGG	GCG	TCG	TTC	TTC	CCG	GAC	GGG	CGC	CGC	ACG	GAC	5764
1896	H	V	H	G	A	P	L	D	R	A	S	F	F	P	D	G	R	R	T	D	1915
5765	CTG	CCC	ACG	TAC	GCC	TTC	CGG	CGC	GAG	CAC	TAC	TGG	CTG	ACG	CCC	GAG	GCC	CGT	ACG	GAC	5824
1916	L	P	T	Y	A	F	R	R	E	H	Y	W	L	T	P	E	A	R	T	D	1935
5825	GCC	CGC	GCA	CTC	GGC	TTC	GAC	CCG	GCG	CGG	CAC	CCG	CTG	CTG	ACG	ACC	ACG	GTC	GAG	GTC	5884
1936	A	R	A	L	G	F	D	P	A	R	H	P	L	L	T	T	T	V	E	V	1955
5885	GCC	GGC	GGC	GAC	GGC	GTC	CTG	CTG	ACC	GGC	CGT	CTC	TCC	CTG	ACC	GAC	CAG	CCC	TGG	CTG	5944
1956	A	G	G	D	G	V	L	L	T	G	R	L	S	L	T	D	Q	P	W	L	1975
5945	GCC	GAC	CAC	ATG	GTC	AAC	GGC	GCC	GTC	CTG	TTG	CCG	GCC	ACC	GCC	TTC	CTG	GAG	CTC	GCC	6004
1976	A	D	H	M	V	N	G	A	V	L	L	P	A	T	A	F	L	E	L	A	1995
6005	CTC	GCG	GCG	GGC	GAC	CAC	GTC	GGG	GCG	GTC	CGG	GTG	GAG	GAA	CTC	ACC	CTC	GAA	GCG	CCG	6064
1996	L	A	A	G	D	H	V	G	A	V	R	V	E	E	L	T	L	E	A	P	2015
6065	CTC	GTC	CTG	CCC	GAG	CGG	GGC	GCC	GTC	CGC	ATC	CAG	GTC	GGC	GTG	AGC	GGC	GAC	GGC	GAG	6124
2016	L	V	L	P	E	R	G	A	V	R	I	Q	V	G	V	S	G	D	G	E	2035

FIG. 23D

6125	TCG	CCG	GCC	GGG	CGC	ACC	TTC	GGT	GTG	TAC	AGC	ACC	CCC	GAC	TCC	GGC	GAC	ACC	GGT	GAC	6184
2036	S	P	A	G	R	T	F	G	V	Y	S	T	P	D	S	G	D	T	G	D	2055
6185	GAC	GCG	CCC	CGG	GAG	TGG	ACC	CGC	CAT	GTC	TCC	GGC	GTA	CTC	GGC	GAA	GGG	GAC	CCG	GCC	6244
2056	D	A	P	R	E	W	T	R	H	V	S	G	V	L	G	E	G	D	P	A	2075
6245	ACG	GAG	TCG	GAC	CAC	CCC	GGC	ACC	GAC	GGG	GAC	GGT	TCA	GCG	GCC	TGG	CCG	CCT	GCG	GCG	6304
2076	T	E	S	D	H	P	G	T	D	G	D	G	S	A	A	W	P	P	A	A	2095
6305	GCG	ACC	GCC	ACA	CCC	CTC	GAC	GGC	GTC	TAC	GAC	CGG	CTC	GCG	GAG	CTC	GGC	TAC	GGA	TAC	6364
2096	A	T	A	T	P	L	D	G	V	Y	D	R	L	A	E	L	G	Y	G	Y	2115
6365	GGT	CCG	GCC	TTC	CAG	GGC	CTG	ACG	GGG	CTG	TGG	CGC	GAC	GGC	GCC	GAC	ACG	CTC	GCC	GAG	6424
2116	G	P	A	F	Q	G	L	T	G	L	W	R	D	G	A	D	T	L	A	E	2135
6425	ATC	CGG	CTG	CCC	GCG	GCG	CAG	CAC	GAG	AGC	GCG	GGG	CTC	TTC	GGC	GTA	CAC	CCG	GCG	CTG	6484
2136	I	R	L	P	A	A	Q	H	E	S	A	G	L	F	G	V	H	P	A	L	2155
6485	CTC	GAC	GCG	GCG	CTC	CAC	CCG	ATC	GTC	CTG	GAG	GGC	AAC	TCA	GCT	GCC	GGT	GCC	TGT	GAC	6544
2156	L	D	A	A	L	H	P	I	V	L	E	G	N	S	A	A	G	A	C	D	2175
6545	GCC	GAT	ACC	GAC	GCG	ACC	GAC	CGG	ATC	CGG	CTG	COG	TTC	GCG	TGG	GCG	GGG	GTG	ACC	CTC	6604
2176	A	D	T	D	A	T	D	R	I	R	L	P	F	A	W	A	G	V	T	L	2195
6605	CAC	GCC	GAA	GGG	GCC	ACC	GCG	CTC	CGC	GTA	CGG	ATC	ACA	CCC	ACC	GGC	CCG	GAC	ACG	GTC	6664
2196	H	A	E	G	A	T	A	L	R	V	R	I	T	P	T	G	P	D	T	V	2215
6665	ACG	CTC	CGC	CTC	ACC	GAC	ACC	ACC	GGT	GCG	CCC	GTG	GCC	ACC	GTG	GAG	TCC	CTG	ACC	CTG	6724
2216	T	L	R	L	T	D	T	T	G	A	P	V	A	T	V	E	S	L	T	L	2235
6725	CGC	GCG	GTG	GCG	AAG	GAC	CGG	CTG	GGC	ACC	ACC	GCC	GGG	CGC	GTC	GAC	GAC	GCC	CTG	TTC	6784
2236	R	A	V	A	K	D	R	L	G	T	T	A	G	R	V	D	D	A	L	F	2255
6785	ACG	GTC	GTG	TGG	ACG	GAG	ACC	GGC	ACA	CCG	GAA	CCC	GCA	GGG	CGC	GGA	GCC	GTG	GAG	GTC	6844
2256	T	V	V	W	T	E	T	G	T	P	E	P	A	G	R	G	A	V	E	V	2275
6845	GAG	GAA	CTC	GTC	GAC	CTC	GCC	GGC	CTC	GGC	GAC	CTC	GTG	GAG	CTC	GGC	GCC	GCG	GAC	GTC	6904
2276	E	E	L	V	D	L	A	G	L	G	D	L	V	E	L	G	A	A	D	V	2295
6905	GTC	CTC	CGG	GCC	GAC	CGC	TGG	ACG	CTC	GAC	GGG	GAC	CCG	TCC	GCC	GCC	GCG	CGC	ACA	GCC	6964
2296	V	L	R	A	D	R	W	T	L	D	G	D	P	S	A	A	A	R	T	A	2315
6965	GTC	CGG	CGC	ACC	CTC	GCC	ATC	GTC	CAG	GAG	TTC	CTG	TCC	GAG	CCG	CGC	TTC	GAC	GGC	TCG	7024
2316	V	R	R	T	L	A	I	V	Q	E	F	L	S	E	P	R	F	D	G	S	2335
7025	CGA	CTG	GTG	TGC	GTC	ACC	AGG	GGC	GCG	GTC	GCC	GCA	CTC	CCC	GGC	GAG	GAC	GTC	ACC	TCC	7084
2336	R	L	V	C	V	T	R	G	A	V	A	A	L	P	G	E	D	V	T	S	2355
7085	CTC	GCC	ACC	GGC	CCC	CTC	TGG	GGC	CTC	GTC	CGC	TCC	GCC	CAG	TCC	GAG	AAC	CCG	GGA	CGC	7144
2356	L	A	T	G	P	L	W	G	L	V	R	S	A	Q	S	E	N	P	G	R	2375
7145	CTG	TTC	CTC	CTG	GAC	CTG	GGT	GAA	GGC	GAA	GGC	GAG	CGC	GAC	GGA	GCC	GAG	GAG	CTG	ATC	7204
2376	L	F	L	L	D	L	G	E	G	E	G	E	R	D	G	A	E	E	L	I	2395
7205	CGC	GCG	GCC	ACG	GCC	GGG	GAC	GAG	CCG	CAG	CTC	GCG	GCA	CCG	GAC	GGC	CGA	CTG	CTC	GCG	7264
2396	R	A	A	T	A	G	D	E	P	Q	L	A	A	R	D	G	R	L	L	A	2415
7265	CCG	AGG	CTG	GCC	CGT	ACC	GCC	GCC	CTT	TGG	AGT	GAG	GAC	ACC	GCC	GGC	GGC	GCC	GAC	CGT	7324
2416	P	R	L	A	R	T	A	A	L	S	S	E	D	T	A	G	G	A	D	R	2435
7325	TTC	GGC	CCC	GAC	GGC	ACC	GTC	CTC	GTC	ACC	GGG	GGC	ACC	GGA	GGC	CTC	GGA	GCG	CTC	CTC	7384
2436	F	G	P	D	G	T	V	L	V	T	G	G	T	G	G	L	G	A	L	L	2455
7385	GCC	CGC	CAC	CTC	GTG	GAG	CGT	CAC	GGG	GTG	CGC	CCG	CTG	CTG	CTG	GTG	AGC	CGC	CGC	GGG	7444
2456	A	R	H	L	V	E	R	H	G	V	R	R	L	L	L	V	S	R	R	G	2475
7445	GCC	GAC	GCC	CCC	GGC	GCG	GCC	GAC	CTG	GGC	GAG	GAC	CTC	GCG	GGC	CTC	GGC	GCG	GAG	GTG	7504
2476	A	D	A	P	G	A	A	D	L	G	E	D	L	A	G	L	G	A	E	V	2495
7505	GCG	TTC	GCC	GCC	GCC	GAC	GCC	GCC	GAC	CGC	GAG	AGC	CTG	GCG	CCG	GCG	ATC	GCC	ACC	GTG	7564
2496	A	F	A	A	A	D	A	A	D	R	E	S	L	A	R	A	I	A	T	V	2515
7565	CCC	GCC	GAG	CAT	CCG	CTG	ACG	GCC	GTC	GTG	CAC	ACG	GCG	GGA	GTC	GTC	GAC	GAC	GCG	ACG	7624
2516	P	A	E	H	P	L	T	A	V	V	H	T	A	G	V	V	D	D	A	T	2535
7625	GTG	GAG	GCG	CTC	ACA	CCG	GAA	CCG	CTG	GAC	GCG	GTA	CTG	CGC	CCG	AAG	GTC	GAC	GCC	GCG	7684
2536	V	E	A	L	T	P	E	R	L	D	A	V	L	R	P	K	V	D	A	A	2555

FIG. 23E

7685	TGG	AAC	CTG	CAC	GAG	CTC	ACC	AAG	GAC	CTG	CGG	CTC	GAC	GCC	TTC	GTC	CTC	TTC	TCC	TCC	7744
2556	W	N	L	H	E	L	T	K	D	L	R	L	D	A	F	V	L	F	S	S	2575
7745	GTC	TCC	GGC	ATC	GTC	GGC	ACC	GCC	GGC	CAG	GCC	AAC	TAC	GCG	GCG	GCC	AAC	ACG	GGC	CTC	7804
2576	V	S	G	I	V	G	T	A	G	Q	A	N	Y	A	A	A	N	T	G	L	2595
7805	GAC	GCC	CTC	GCC	GCC	CAC	CGC	GCC	GCC	ACG	GGC	CTG	GCC	GCC	ACG	TCG	CTG	GCC	TGG	GGC	7864
2596	D	A	L	A	A	H	R	A	A	T	G	L	A	A	T	S	L	A	W	G	2615
7865	CTC	TGG	GAC	GGC	ACG	CAC	GGC	ATG	GGC	GGC	ACG	CTC	GGC	GCC	GCC	GAC	CTC	GCC	CGC	TGG	7924
2616	L	W	D	G	T	H	G	M	G	G	T	L	G	A	A	D	L	A	R	W	2635
7925	AGC	CGG	GCC	GGA	ATC	ACC	CCG	CTC	ACC	CCG	CTG	CAG	GGC	CTC	GCG	CTC	TTC	GAC	GCC	GCG	7984
2636	S	R	A	G	I	T	P	L	T	P	L	Q	G	L	A	L	F	D	A	A	2655
7985	GTC	GCC	AGG	GAC	GAC	GCC	CTC	CTC	GTA	CCC	GCC	GGG	CTC	CGT	CCC	ACC	GCC	CAC	CGG	GGC	8044
2656	V	A	R	D	D	A	L	L	V	P	A	G	L	R	P	T	A	H	R	G	2675
8045	ACG	GAC	GGA	CAG	CCT	CCT	GCG	CTG	TGG	CGC	GGC	CTC	GTC	CGG	GCG	CGC	CCG	CGC	CGT	GCC	8104
2676	T	D	G	Q	P	P	A	L	W	R	G	L	V	R	A	R	P	R	R	A	2695
8105	GCG	CGG	ACG	GCC	GCC	GAG	GCG	GCG	GAC	ACG	ACC	GGC	GGC	TGG	CTG	AGC	GGG	CTC	GCC	GCA	8164
2696	A	R	T	A	A	E	A	A	D	T	T	G	G	W	L	S	G	L	A	A	2715
8165	CAG	TCC	CCC	GAG	GAG	CGG	CGC	AGC	ACA	GCC	GTC	ACG	CTC	GTC	ACG	GGT	GTC	GTC	GCG	GAC	8224
2716	Q	S	P	E	E	R	R	S	T	A	V	T	L	V	T	G	V	V	A	D	2735
8225	GTC	CTC	GGG	CAC	GCC	GAC	TCC	GCC	GCG	GTC	GGG	GCG	GAG	CGG	TCC	TTC	AAG	GAC	CTC	GGC	8284
2736	V	L	G	H	A	D	S	A	A	V	G	A	E	R	S	F	K	D	L	G	2755
8285	TTC	GAC	TCC	CTG	GCC	GGG	GTC	GAG	CTC	CGC	AAC	CGG	CTG	AAC	GCC	GCC	ACC	GGC	CTG	CGG	8344
2756	F	D	S	L	A	G	V	E	L	R	N	R	L	N	A	A	T	G	L	R	2775
8345	CTC	CCC	GCG	ACC	ACG	GTC	TTC	GAC	CAT	CCC	TCG	CGG	GCC	GCG	CTC	GCG	TCC	CAT	CTC	CTC	8404
2776	L	P	A	T	T	V	F	D	H	P	S	P	A	A	L	A	S	H	L	L	2795
8405	GCC	CAG	GTC	CCC	GGG	TTG	AAG	GAG	GGG	ACG	GCG	GCG	ACC	GCG	ACC	GTC	GTC	GCC	GAG	CGG	8464
2796	A	Q	V	P	G	L	K	E	G	T	A	A	T	A	T	V	V	A	E	R	2815
8465	GGC	GCT	TCC	TTC	GGT	GAC	CGT	GCG	ACC	GAC	GAC	GAT	CCG	ATC	GCG	ATC	GTC	GGC	ATG	GCA	8524
2816	G	A	S	F	G	D	R	A	T	D	D	D	P	I	A	I	V	G	M	A	2835
8525	TGC	CGC	TAT	CCG	GGT	GGT	GTC	TCG	TCG	CCG	GAG	GAC	CTG	TGG	CGG	CTG	GTC	GCC	GAG	GGG	8584
2836	C	R	Y	P	G	G	V	S	S	P	E	D	L	W	R	L	V	A	E	G	2855
8585	ACG	GAC	GCG	ATC	AGC	GAG	TTC	CCC	GTC	AAC	CGC	GGC	TGG	GAC	CTG	GAG	AGC	CTC	TAC	GAC	8644
2856	T	D	A	I	S	E	F	P	V	N	R	G	W	D	L	E	S	L	Y	D	2875
8645	CCG	GAT	CCC	GAG	TCG	AAG	GGC	ACC	ACG	TAC	TGC	CGG	GAG	GGC	GGG	TTC	CTG	GAA	GGC	GCC	8704
2876	P	D	P	E	S	K	G	T	T	Y	C	R	E	G	G	F	L	E	G	A	2895
8705	GGT	GAC	TTC	GAC	GCC	GCC	TTC	TTC	GGC	ATC	TCG	CCG	CGC	GAG	GCC	CTG	GTC	ATG	GAC	CCG	8764
2896	G	D	F	D	A	A	F	F	G	I	S	P	R	E	A	L	V	M	D	P	2915
8765	CAG	CAG	CGG	CTG	CTG	CTG	GAG	GTC	TCC	TGG	GAG	GCG	CTG	GAA	CGC	GCG	GGC	ATC	GAC	CCG	8824
2916	Q	Q	R	L	L	L	E	V	S	W	E	A	L	E	R	A	G	I	D	P	2935
8825	TCC	TCG	CTG	CGC	GGC	AGC	CGC	GGT	GGT	GTC	TAC	GTC	GGC	GCC	GCG	CAC	GGC	TCG	TAC	GCC	8884
2936	S	S	L	R	G	S	R	G	G	V	Y	V	G	A	A	H	G	S	Y	A	2955
8885	TCC	GAT	CCC	CGG	CTG	GTC	CCC	GAG	GGC	TCG	GAG	GGC	TAT	CTG	CTG	ACC	GGC	AGC	GCC	GAC	8944
2956	S	D	P	R	L	V	P	E	G	S	E	G	Y	L	L	T	G	S	A	D	2975
8945	GCG	GTC	ATG	TCC	GGC	CGC	ATC	TCC	TAC	GCG	CTC	GGT	CTC	GAA	GGA	CCG	TCC	ATG	ACG	GTC	9004
2976	A	V	M	S	G	R	I	S	Y	A	L	G	L	E	G	P	S	M	T	V	2995
9005	GAG	ACG	GCC	TGC	TCC	TCC	TCG	CTG	GTC	GCG	CTG	CAT	CTG	GCG	GTA	CGG	GCG	CTG	CGG	CAC	9064
2996	E	T	A	C	S	S	S	L	V	A	L	H	L	A	V	R	A	L	R	H	3015
9065	GGC	GAG	TGC	GGG	CTC	GCG	CTG	GCG	GGC	GGG	GTC	GCG	GTC	ATG	GCC	GAT	CCG	GCG	GCG	TTC	9124
3016	G	E	C	G	L	A	L	A	G	G	V	A	V	M	A	D	P	A	A	F	3035
9125	GTC	GAG	TTC	TCC	CGG	CAG	AAG	GGG	CTG	GCC	GCC	GAC	GGC	CGC	TGC	AAG	GCG	TTC	TCG	GCC	9184
3036	V	E	F	S	R	Q	K	G	L	A	A	D	G	R	C	K	A	F	S	A	3055
9185	GCC	GCC	GAC	GGC	ACC	GGC	TGG	GCC	GAG	GGC	GTC	GGC	GTC	GTC	CTG	GAG	CGG	CTG	TCG		9244
3056	A	A	D	G	T	G	W	A	E	G	V	G	V	L	V	L	E	R	L	S	3075

FIG. 23F

9245	GAC	GCG	CGC	CGC	GCG	GGG	CAC	ACG	GTC	CTC	GGC	CTG	GTC	ACC	GGC	ACC	GCG	GTC	AAC	CAG	9304
3076	D	A	R	R	A	G	H	T	V	L	G	L	V	T	G	T	A	V	N	Q	3095
9305	GAC	GGT	GCC	TCC	AAC	GGG	CTG	ACC	GCG	CCC	AAC	GGC	CCA	GCC	CAG	CAA	CGC	GTC	ATC	GCC	9364
3096	D	G	A	S	N	G	L	T	A	P	N	G	P	A	Q	Q	R	V	I	A	3115
9365	GAG	GCG	CTC	GCC	GAC	GCC	GGG	CTG	TCC	CCG	GAG	GAC	GTG	GAC	GCG	GTC	GAG	GCG	CAC	GGC	9424
3116	E	A	L	A	D	A	G	L	S	P	E	D	V	D	A	V	E	A	H	G	3135
9425	ACC	GGC	ACC	CGG	CTC	GGC	GAC	CCC	ATC	GAG	GCC	GGG	GCG	CTG	CTC	GCC	GCC	TCC	GGA	CGG	9484
3136	T	G	T	R	L	G	D	P	I	E	A	G	A	L	L	A	A	S	G	R	3155
9485	AAC	CGT	TCC	GGC	GAC	CAC	CCG	CTG	TGG	CTC	GGC	TCG	CTG	AAG	TCC	AAC	ATC	GGG	CAT	GCC	9544
3156	N	R	S	G	D	H	P	L	W	L	G	S	L	K	S	N	I	G	H	A	3175
9545	CAG	GCC	GCC	GCC	GGT	GTC	GGC	GGC	GTC	ATC	AAG	ATG	CTC	CAG	GCG	CTG	CGG	CAC	GGC	TTG	9604
3176	Q	A	A	A	G	V	G	G	V	I	K	M	L	Q	A	L	R	H	G	L	3195
9605	CTG	CCC	CGC	ACC	CTC	CAC	GCC	GAC	GAG	CCG	ACC	CCG	CAT	GCC	GAC	TGG	AGC	TCC	GGC	CGG	9664
3196	L	P	R	T	L	H	A	D	E	P	T	P	H	A	D	W	S	S	G	R	3215
9665	GTA	CGG	CTG	CTC	ACC	TCC	GAG	GTG	CCG	TGG	CAG	CGG	ACC	GGC	CGG	CCC	CGG	CGG	ACC	GGG	9724
3216	V	R	L	L	T	S	E	V	P	W	Q	R	T	G	R	P	R	R	T	G	3235
9725	GTG	TCC	GCC	TTC	GGC	GTC	GGC	GGC	ACC	AAT	GCC	CAT	GTC	GTG	CTC	GAA	GAG	GCA	CCC	GCC	9784
3236	V	S	A	F	G	V	G	G	T	N	A	H	V	V	L	E	E	A	P	A	3255
9785	CCG	CCC	GCG	CCG	GAA	CCG	GCC	GGG	GAG	GCC	CCC	GGC	GGC	TCC	CGC	GCC	GCA	GAA	GGG	GCG	9844
3256	P	P	A	P	E	P	A	G	E	A	P	G	G	S	R	A	A	E	G	A	3275
9845	GAA	GGG	CCC	CTG	GCC	TGG	GTG	GTC	TCC	GGA	CGC	GAC	GAG	CCG	GCC	CTG	CGG	TCC	CAG	GCC	9904
3276	E	G	P	L	A	W	V	V	S	G	R	D	E	P	A	L	R	S	Q	A	3295
9905	CGG	CGG	CTC	CGC	GAC	CAC	CTC	TCC	CGC	ACC	CCC	GGG	GCC	CGC	CCG	CGT	GAC	ATC	GCC	TTC	9964
3296	R	R	L	R	D	H	L	S	R	T	P	G	A	R	P	R	D	I	A	F	3315
9965	TCC	CTC	GCC	GCC	ACG	CGC	GCA	GCC	TTT	GAC	CAC	CGC	GCC	GTG	CTG	ATC	GGC	TCG	GAC	GGG	10024
3316	S	L	A	A	T	R	A	A	F	D	H	R	A	V	L	I	G	S	D	G	3335
10025	GCC	GAA	CTC	GCC	GCC	GCC	CTG	GAC	GCG	TTG	GCC	GAA	GGA	CGC	GAC	GGT	CCG	GCG	GTG	GTG	10084
3336	A	E	L	A	A	A	L	D	A	L	A	E	G	R	D	G	P	A	V	V	3355
10085	CGC	GGA	GTC	CGC	GAC	CGG	GAC	GGC	AGG	ATG	GCC	TTC	CTC	TTC	ACC	GGG	CAG	GGC	AGC	CAG	10144
3356	R	G	V	R	D	R	D	G	R	M	A	F	L	F	T	G	Q	G	S	Q	3375
10145	CGC	GCC	GGG	ATG	GCC	CAC	GAC	CTG	CAT	GCC	GCC	CAT	ACC	TTC	TTC	GCG	TCC	GCC	CTC	GAC	10204
3376	R	A	G	M	A	H	D	L	H	A	A	H	T	F	F	A	S	A	L	D	3395
10205	GAG	GTG	ACG	GAC	CGT	CTC	GAC	CCG	CTG	CTC	GGC	CGG	CCG	CTC	GGC	GCG	CTG	CTG	GAC	GCC	10264
3396	E	V	T	D	R	L	D	P	L	L	G	R	P	L	G	A	L	L	D	A	3415
10265	CGA	CCC	GGC	TCG	CCC	GAA	GCG	GCA	CTC	CTG	GAC	CGG	ACC	GAG	TAC	ACC	CAG	CCG	GCG	CTC	10324
3416	R	P	G	S	P	E	A	A	L	L	D	R	T	E	Y	T	Q	P	A	L	3435
10325	TTC	GCC	GTC	GAG	GTG	GCG	CTC	CAC	CGG	CTG	CTG	GAG	CAC	TGG	GGG	ATG	CGC	CCC	GAC	CTG	10384
3436	F	A	V	E	V	A	L	H	R	L	L	E	H	W	G	M	R	P	D	L	3455
10385	CTG	CTG	GGG	CAC	TCG	GTG	GGC	GAA	CTG	GCG	GCC	GCC	CAC	GTC	GCG	GGT	GTG	CTC	GAT	CTC	10444
3456	L	L	G	H	S	V	G	E	L	A	A	A	H	V	A	G	V	L	D	L	3475
10445	CAC	GAC	GCC	TGC	GCG	CTG	GTG	GCC	GCC	CGC	GGC	AGG	CTG	ATG	CAG	CGC	CTG	CCG	CCC	GGC	10504
3476	D	D	A	C	A	L	V	A	A	R	G	R	L	M	Q	R	L	P	P	G	3495
10505	GGC	GCG	ATG	GTC	TCC	GTG	CGG	GCC	GGC	GAG	GAC	GAG	GTC	CGC	GCA	CTG	CTG	GCC	GGC	CGC	10564
3496	G	A	M	V	S	V	R	A	G	E	D	E	V	R	A	L	L	A	G	R	3515
10565	GAG	GAC	GCC	GTC	TGC	GTG	GCC	GCG	GTG	AAC	GGC	CCC	CGG	TCG	GTG	GTG	ATC	TCC	GGC	GCG	10624
3516	E	D	A	V	C	V	A	A	V	N	G	P	R	S	V	V	I	S	G	A	3535
10625	GAG	GAA	GCG	GTG	GCC	GAG	GCG	GCG	GCG	CAG	CTC	GCC	GGA	CGA	GGC	CGC	CGC	ACC	AGG	CGG	10684
3536	E	E	A	V	A	E	A	A	A	Q	L	A	G	R	G	R	R	T	R	R	3555
10685	CTC	CGC	GTC	GCG	CAC	GCC	TTC	CAC	TCA	CCC	CTG	ATG	GAC	GGC	ATG	CTC	GCC	GGA	TTC	CGG	10744
3556	L	R	V	A	H	A	F	H	S	P	L	M	D	G	M	L	A	G	F	R	3575
10745	GAG	GTC	GCC	GCC	GGC	CTG	CGC	TAC	CGG	GAA	CCG	GAG	CTG	ACG	GTC	GTC	TCC	ACG	GTC	ACG	10804
3576	E	V	A	A	G	L	R	Y	R	E	P	E	L	T	V	V	S	T	V	T	3595

FIG. 23G

10805	GGG	CGG	CCC	GCC	CGC	CCC	GGT	GAA	CTC	ACC	GGC	CCC	GAC	TAC	TGG	GTG	GCC	CAG	GTC	CGT	10864
3596	G	R	P	A	R	P	G	E	L	T	G	P	D	Y	W	V	A	Q	V	R	3615
10865	GAG	CCC	GTG	CGC	TTC	GCG	GAC	GCG	GTC	CGC	ACG	GCA	CAC	CGC	CTC	GGA	GCC	CGC	ACC	TTC	10924
3616	E	P	V	R	F	A	D	A	V	R	T	A	H	R	L	G	A	R	T	F	3635
10925	CTG	GAG	ACC	GGC	CCG	GAC	GGC	GTG	CTG	TGC	GGC	ATG	GCA	GAG	GAG	TGC	CTG	GAG	GAC	GAC	10984
3636	L	E	T	G	P	D	G	V	L	C	G	M	A	E	E	C	L	E	D	D	3655
10985	ACC	GTG	GCC	CTG	CTG	CCG	GCG	ATC	CAC	AAG	CCC	GGC	ACC	GCG	CCG	CAC	GGT	CCG	GCG	GCT	11044
3656	T	V	A	L	L	P	A	I	H	K	P	G	T	A	P	H	G	P	A	A	3675
11045	CCC	GGC	GCG	CTG	CGG	GCG	GCC	GCC	GCC	GCG	TAC	GGC	CGG	GGC	GCC	CGG	GTG	GAC	TGG	GCC	11104
3676	P	G	A	L	R	A	A	A	A	A	Y	G	R	G	A	R	V	D	W	A	3695
11105	GGG	ATG	CAC	GCC	GAC	GGC	CCC	GAG	GGG	CCG	GCC	CGC	CGC	GTG	GAA	CTG	CCC	GTC	CAC	GCC	11164
3696	G	M	H	A	D	G	P	E	G	P	A	R	R	V	E	L	P	V	H	A	3715
11165	TTC	CGG	CAC	CGC	CGC	TAC	TGG	CTC	GCC	CCG	GGC	CGC	GCG	GCG	GAC	ACC	GAC	GAC	TGG	ATG	11224
3716	F	R	H	R	R	Y	W	L	A	P	G	R	A	A	D	T	D	D	W	M	3735
11225	TAC	CGG	ATC	GGC	TGG	GAC	CGG	CTG	CCG	GCT	GTG	ACC	GGC	GGG	GCC	CGG	ACC	GCC	GGC	CGC	11284
3736	Y	R	I	G	W	D	R	L	P	A	V	T	G	G	A	R	T	A	G	R	3755
11285	TGG	CTG	GTG	ATC	CAC	CCC	GAC	AGC	CCG	CGC	TGC	CGG	GAG	CTG	TCC	GGC	CAC	GCC	GAA	CGC	11344
3756	W	L	V	I	H	P	D	S	P	R	C	R	E	L	S	G	H	A	E	R	3775
11345	GCG	CTG	CGC	GCC	GCG	GGC	GCG	AGC	CCC	GTA	CCG	CTG	CCC	GTG	GAC	GCT	CCG	GCC	GCC	GAC	11404
3776	A	L	R	A	A	G	A	S	P	V	P	L	P	V	D	A	P	A	A	D	3795
11405	CGG	GCG	TCC	TTC	GCG	GCA	CTG	CTG	CGC	TCC	GCC	ACC	GGA	CCT	GAC	ACA	CGA	GGT	GAC	ACA	11464
3796	R	A	S	F	A	A	L	L	R	S	A	T	G	P	D	T	R	G	D	T	3815
11465	GCC	GCG	CCC	GTG	GCC	GGT	GTG	CTG	TGG	CTG	CTG	TCC	GAG	GAG	GAT	CGG	CCC	CAT	CGC	CAG	11524
3816	A	A	P	V	A	G	V	L	S	L	L	S	E	E	D	R	P	H	R	Q	3835
11525	CAC	GCC	CCG	GTA	CCC	GCC	GGG	GTG	CTG	GCG	ACG	CTG	TCC	CTG	ATG	CAG	GCT	ATG	GAG	GAG	11584
3836	H	A	P	V	P	A	G	V	L	A	T	L	S	L	M	Q	A	M	E	E	3855
11585	GAG	GCG	GTG	GAG	GCT	CGC	GTG	TGG	TGC	GTG	TCC	CGC	GCC	GCG	GTG	GCC	GCC	GCC	GAC	CGG	11644
3856	E	A	V	E	A	R	V	W	C	V	S	R	A	A	V	A	A	A	D	R	3875
11645	GAA	CGG	CCC	GTG	GGC	GCG	GGC	GCC	GCC	CTG	TGG	GGG	CTG	GGG	CGG	GTG	GCC	GCC	CTG	GAA	11704
3876	E	R	P	V	G	A	G	A	A	L	W	G	L	G	R	V	A	A	L	E	3895
11705	CGC	CCC	ACC	CGG	TGG	GGC	GGT	CTC	GTG	GAC	CTG	CCC	GCC	TGG	CCC	GGT	GCG	GCG	CAC	TGG	11764
3896	R	P	T	R	W	G	G	L	V	D	L	P	A	S	P	G	A	A	H	W	3915
11765	GCG	GCC	GCC	GTG	GAA	CGG	CTC	GCC	GGT	CCC	GAG	GAC	CAG	ATC	GCC	GTG	CGC	GCG	TCC	GGC	11824
3916	A	A	A	V	E	R	L	A	G	P	E	D	Q	I	A	V	R	A	S	G	3935
11825	AGT	TGG	GGC	CGG	CGC	CTC	ACC	AGG	CTG	CCG	CGC	GAC	GGC	GGC	GGC	CGG	ACG	GCC	GCA	CCC	11884
3936	S	W	G	R	R	L	T	R	L	P	R	D	G	G	G	R	T	A	A	P	3955
11885	GCG	TAC	CGG	CCG	CGC	GGC	ACG	GTG	CTC	GTG	ACC	GGT	GGC	ACC	GGC	GCG	CTC	GGC	GGG	CAT	11944
3956	A	Y	R	P	R	G	T	V	L	V	T	G	G	T	G	A	L	G	G	H	3975
11945	CTC	GCC	CGC	TGG	CTC	GCC	GCG	GCG	GGC	GCC	GAA	CAC	CTG	GCG	CTC	ACC	AGC	CGC	CGG	GGC	12004
3976	L	A	R	W	L	A	A	A	G	A	E	H	L	A	L	T	S	R	R	G	3995
12005	CCG	GAC	GCG	CCC	GGC	GCC	GCC	GGA	CTC	GAG	GCC	GAA	CTC	CTC	CTC	CTG	GCC	GCC	AAG	GTG	12064
3996	P	D	A	P	G	A	A	G	L	E	A	E	L	L	L	L	G	A	K	V	4015
12065	ACG	TTC	GCC	GCC	TGC	GAC	ACC	GCC	GAC	CGC	GAC	GGC	CTC	GCC	CGG	GTG	CTG	CGG	GCG	ATA	12124
4016	T	F	A	A	C	D	T	A	D	R	D	G	L	A	R	V	L	R	A	I	4035
12125	CCG	GAG	GAC	ACC	CCG	CTC	ACC	GCG	GTG	TTC	CAC	GCC	GCG	GGC	GTA	CCG	CAG	GTC	ACG	CCG	12184
4036	P	E	D	T	P	L	T	A	V	F	H	A	A	G	V	P	Q	V	T	P	4055
12185	CTG	TCC	CGT	ACC	TGG	CCC	GAG	CAC	TTC	GCC	GAC	GTG	TAC	GCG	GGC	AAG	GCG	GCG	GGC	GCC	12244
4056	L	S	R	T	S	P	E	H	F	A	D	V	Y	A	G	K	A	A	G	A	4075
12245	GCG	CAC	CTG	GAC	GAA	CTG	ACC	CGC	GAA	CTC	GGC	GCC	GGA	CTC	GAC	GCG	TTC	GTC	CTC	TAC	12304
4076	A	H	L	D	E	L	T	R	E	L	G	A	G	L	D	A	F	V	L	Y	4095
12305	TCC	TCC	GGC	GCC	GGC	GTG	TGG	GGC	AGC	GCC	GGC	CAG	GGT	GCC	TAC	GCC	GCC	GCC	AAC	GCC	12364
4096	S	S	G	A	G	V	W	G	S	A	G	Q	G	A	Y	A	A	A	N	A	4115

FIG. 23H

12365	GCC	CTG	GAC	GCG	CTC	GCC	CGG	CGC	CGT	GCG	GCG	GAC	GGA	CTC	CCC	GCC	ACC	TCC	ATC	GCC	12424
4116	A	L	D	A	L	A	R	R	R	A	A	D	G	L	P	A	T	S	I	A	4135
12425	TGG	GGC	GTG	TGG	GGC	GGC	GGC	GGT	ATG	GGG	GCC	GAC	GAG	GCG	GGC	GCG	GAG	TAT	CTG	GGC	12484
4136	W	G	V	W	G	G	G	G	M	G	A	D	E	A	G	A	E	Y	L	G	4155
12485	CGG	CGC	GGT	ATG	CGC	CCC	ATG	GCA	CGG	GTC	TCC	GCG	CTC	CGG	GCG	ATG	GCC	ACC	GCC	ATC	12544
4156	R	R	G	M	R	P	M	A	P	V	S	A	L	R	A	M	A	T	A	I	4175
12545	GCC	TCC	GGG	GAA	CCC	TGC	CCC	ACC	GTC	ACC	CAC	ACC	GAC	TGG	GAG	CGC	TTC	GGC	GAG	GGC	12604
4176	A	S	G	E	P	C	P	T	V	T	H	T	D	W	E	R	F	G	E	G	4195
12605	TTC	ACC	GCC	TTC	CGG	CCC	AGC	CCT	CTG	ATC	GCG	GGG	CTC	GGC	ACG	CCG	GGC	GGC	GGC	CGG	12664
4196	F	T	A	F	R	P	S	P	L	I	A	G	L	G	T	P	G	G	G	R	4215
12665	GCG	GCG	GAG	ACC	CCC	GAG	GAG	GGG	AAC	GCC	ACC	GCT	GCG	GCG	GAC	CTC	ACC	GCC	CTG	CCG	12724
4216	A	A	E	T	P	E	E	G	N	A	T	A	A	A	D	L	T	A	L	P	4235
12725	CCC	GCC	GAA	CTC	CGC	ACC	GCG	CTG	CGC	GAG	CTG	GTG	CGA	GCC	CGG	ACC	GCC	GCG	GCG	CTC	12784
4236	P	A	E	L	R	T	A	L	R	E	L	V	R	A	R	T	A	A	A	L	4255
12785	GGC	CTC	GAC	GAC	CGG	GCC	GAG	GTC	GCC	GAG	GGC	GAA	CGG	TTC	CCC	GCC	ATG	GGC	TTC	GAC	12844
4256	G	L	D	D	P	A	E	V	A	E	G	E	R	F	P	A	M	G	F	D	4275
12845	TCC	CTG	GCC	ACC	GTA	CGG	CTG	CGC	CGC	GGA	CTC	GCC	TGG	GCC	ACG	GGC	CTC	GAC	CTG	CCC	12904
4276	S	L	A	T	V	R	L	R	R	G	L	A	S	A	T	G	L	D	L	P	4295
12905	CCC	GAT	CTG	CTC	TTC	GAC	CGG	GAC	ACC	CCG	GCC	GCG	CTC	GCC	GCC	CAC	CTG	GCC	GAA	CTG	12964
4296	P	D	L	L	F	D	R	D	T	P	A	A	L	A	A	H	L	A	E	L	4315
12965	CTC	GCC	ACC	GCA	CGG	GAC	CAC	GGA	CCC	GGC	GGC	CCC	GGG	ACC	GGT	GCC	GCG	CCG	GCC	GAT	13024
4316	L	A	T	A	R	D	H	G	P	G	G	P	G	T	G	A	A	P	A	D	4335
13025	GCC	GGA	AGC	GGC	CTG	CCG	GCC	CTC	TAC	CGG	GAG	GCC	GTC	CGC	ACC	GGC	CGG	GCC	GCG	GAA	13084
4336	A	G	S	G	L	P	A	L	Y	R	E	A	V	R	T	G	R	A	A	E	4355
13085	ATG	GCC	GAA	CTG	CTC	GCC	GCC	GCT	TCC	CGG	TTC	CGC	CCC	GCC	TTC	GGG	ACG	GCG	GAC	CGG	13144
4356	M	A	E	L	L	A	A	A	S	R	F	R	P	A	F	G	T	A	D	R	4375
13145	CAG	CCG	GTG	GCC	CTC	GTG	CCG	CTG	GCC	GAC	GGC	GCG	GAG	GAC	ACC	GGG	CTC	CCG	CTG	CTC	13204
4376	Q	P	V	A	L	V	P	L	A	D	G	A	E	D	T	G	L	P	L	L	4395
13205	GTG	GGC	TGC	GCC	GGG	ACG	GCG	GTG	GCC	TCC	GGC	CCG	GTG	GAG	TTC	ACC	GCC	TTC	GCC	GGA	13264
4396	V	G	C	A	G	T	A	V	A	S	G	P	V	E	F	T	A	F	A	G	4415
13265	GCG	CTG	GCG	GAC	CTC	COG	GCG	GCG	GCC	COG	ATG	GCC	GCG	CTG	CCG	CAG	CCC	GGC	TTT	CTG	13324
4416	A	L	A	D	L	P	A	A	A	P	M	A	A	L	P	Q	P	G	F	L	4435
13325	CCG	GGA	GAA	CGA	GTC	CCG	GCC	ACC	COG	GAG	GCA	TTG	TTC	GAG	GCC	CAG	GCG	GAA	GCG	CTG	13384
4436	P	G	E	R	V	P	A	T	P	E	A	L	F	E	A	Q	A	E	A	L	4455
13385	CTG	CGC	TAC	GCG	GCC	GGC	CGG	CCC	TTC	GTG	CTG	CTG	GGG	CAC	TCC	GCC	GGC	GCC	AAC	ATG	13444
4456	L	R	Y	A	A	G	R	P	F	V	L	L	G	H	S	A	G	A	N	M	4475
13445	GCC	CAC	GCC	CTG	ACC	CGT	CAT	CTG	GAG	GCG	AAC	GGT	GGC	GGC	CCC	GCA	GGG	CTG	GTG	CTC	13504
4476	A	H	A	L	T	R	H	L	E	A	N	G	G	G	P	A	G	L	V	L	4495
13505	ATG	GAC	ATC	TAC	ACC	CCC	GCC	GAC	CCC	GGC	GCG	ATG	GGC	GTC	TGG	CGG	AAC	GAC	ATG	TTC	13564
4496	M	D	I	Y	T	P	A	D	P	G	A	M	G	V	W	R	N	D	M	F	4515
13565	CAG	TGC	GTC	TGG	CGG	CGC	TGC	GAC	ATC	CCC	CCG	GAC	GAC	CAC	CGC	CTC	ACG	GCC	ATG	GGC	13624
4516	Q	W	V	W	R	R	S	D	I	P	P	D	D	H	R	L	T	A	M	G	4535
13625	GCC	TAC	CAC	CGG	CTG	CTT	CTC	GAC	TGG	TGC	CCC	ACC	CCC	GTC	CGC	GCC	CCC	GTA	CTG	CAT	13684
4536	A	Y	H	R	L	L	L	D	W	S	P	T	P	V	R	A	P	V	L	H	4555
13685	CTG	CGC	GCC	GCG	GAA	CCC	ATG	GGC	GAC	TGG	CCA	CCC	GGG	GAC	ACC	GGC	TGG	CAG	TCC	CAC	13744
4556	L	R	A	A	E	P	M	G	D	W	P	P	G	D	T	G	W	Q	S	H	4575
13745	TGG	GAC	GGC	GCG	CAC	ACC	ACC	GCC	GGC	ATC	CCC	GGA	AAC	CAC	TTC	ACG	ATG	ATG	ACC	GAA	13804
4576	W	D	G	A	H	T	T	A	G	I	P	G	N	H	F	T	M	M	T	E	4595
13805	CAC	GCC	TCC	GCC	GCC	GCC	CGG	CTC	GTG	CAC	GGC	TGG	CTC	GCG	GAA	CGG	ACC	CCG	TCC	GGG	13864
4596	H	A	S	A	A	A	R	L	V	H	G	W	L	A	E	R	T	P	S	G	4615
13865	CAG	GGC	GGG	TCA	CCG	TCC	CGC	GCG	GCG	GGG	AGA	GAG	GAG	AGG	CCG	TGA	ACACGGCAGCCGGCCCC				13928
4616	Q	G	G	S	P	S	R	A	A	G	R	E	E	R	P	*					4631

FIG. 231

13929 GACCGGCACCGCCGCGCGCGGCACACCGCCCGCGCGCGGCACACGACCTGTCCCGCGCCGACGAGGCTCCAACTCA 14008

14009 CCGGGCCGCACAGTGGTTTCGCGCGCAACCAGGGAGACCCCTACGGG ATG ATC CTG CGC GCC GGC ACC GCC 14079
1 M I L R A G T A 8

14080 GAC CGG GCA CCG TAC GAG GAA GAG ATC CCC GGG TAC CGA GCT CGA ATT CTT AAT TAA GGAG 14140
9 D P A P Y E E E I P G Y R A R I L N * 27

14141 GTCGTAG ATG AGT AAC AAG AAC AAC GAT GAG CTG CAG CGG CAG GCC TCG GAA AAC ACC CTG 14201
1 M S N K N N D E L Q R Q A S E N T L 18

14202 GGG CTG AAC CCG GTC ATC GGT ATC CGC CGC AAA GAC CTG TTG AGC TCG GCA CGC ACC GTG 14261
19 G L N P V I G I R R K D L L S S A R T V 38

14262 CTG CGC CAG GCC GTG CGC CAA CCG CTG CAC AGC GCC AAG CAT GTG GCC CAC TTT GGC CTG 14321
39 L R Q A V R Q P L H S A K H V A H F G L 58

14322 GAG CTG AAG AAC GTG CTG CTG GGC AAG TCC AGC CTT GCC CCG GAA AGC GAC GAC CGT CGC 14381
59 E L K N V L L G K S S L A P E S D D R R 78

14382 TTC AAT GAC CCG GCA TGG AGC AAC AAC CCA CTT TAC CGC CGC TAC CTG CAA ACC TAT CTG 14441
79 F N D P A W S N N P L Y R R Y L Q T Y L 98

14442 GCC TGG CGC AAG GAG CTG CAG GAC TGG ATC GGC AAC AGC GAC CTG TCG CCC CAG GAC ATC 14501
99 A W R K E L Q D W I G N S D L S P Q D I 118

14502 AGC CGC GGC CAG TTC GTC ATC AAC CTG ATG ACC GAA GCC ATG GCT CCG ACC AAC ACC CTG 14561
119 S R G Q F V I N L M T E A M A P T N T L 138

14562 TCC AAC CCG GCA GCA GTC AAA CGC TTC TTC GAA ACC GGC GGC AAG AGC CTG CTC GAT GGC 14621
139 S N P A A V K R F F E T G G K S L L D G 158

14622 CTG TCC AAC CTG GCC AAG GAC CTG GTC AAC AAC GGT GGC ATG CCC AGC CAG GTG AAC ATG 14681
159 L S N L A K D L V N N G G M P S Q V N M 178

14682 GAC GCC TTC GAG GTG GGC AAG AAC CTG GGC ACC AGT GAA GGC GCC GTG GTG TAC CGC AAC 14741
179 D A F E V G K N L G T S E G A V V Y R N 198

14742 GAT GTG CTG GAG CTG ATC CAG TAC AAG CCC ATC ACC GAG CAG GTG CAT GCC CGC CCG CTG 14801
199 D V L E L I Q Y K P I T E Q V H A R P L 218

14802 CTG GTG GTG CCG CCG CAG ATC AAC AAG TTC TAC GTA TTC GAC CTG AGC CCG GAA AAG AGC 14861
219 L V V P P Q I N K F Y V F D L S P E K S 238

14862 CTG GCA CGC TAC TGC CTG CGC TCG CAG CAG CAG ACC TTC ATC ATC AGC TGG CGC AAC CCG 14921
239 L A R Y C L R S Q Q Q T F I I S W R N P 258

14922 ACC AAA GCC CAG CGC GAA TGG GGC CTG TCC ACC TAC ATC GAC GCG CTC AAG GAG GCG GTC 14981
259 T K A Q R E W G L S T Y I D A L K E A V 278

14982 GAC GCG GTG CTG GCG ATT ACC GGC AGC AAG GAC CTG AAC ATG CTC GGT GCC TGC TCC GGC 15041
279 D A V L A I T G S K D L N M L G A C S G 298

15042 GGC ATC ACC TGC ACG GCA TTG GTC GGC CAC TAT GCC GCC CTC GGC GAA AAC AAG GTC AAT 15101
299 G I T C T A L V G H Y A A L G E N K V N 318

15102 GCC CTG ACC CTG CTG GTC AGC GTG CTG GAC ACC ACC ATG GAC AAC CAG GTC GCC CTG TTC 15161
319 A L T L L V S V L D T T M D N Q V A L F 338

15162 GTC GAC GAG CAG ACT TTG GAG GCC GCC AAG CGC CAC TCC TAC CAG GCC GGT GTG CTC GAA 15221
339 V D E Q T L E A A K R H S Y Q A G V L E 358

15222 GGC AGC GAG ATG GCC AAG GTG TTC GCC TGG ATG CGC CCC AAC GAC CTG ATC TGG AAC TAC 15281
359 G S E M A K V F A W M R P N D L I W N Y 378

15282 TGG GTC AAC AAC TAC CTG CTC GGC AAC GAG CCG CCG GTG TTC GAC ATC CTG TTC TGG AAC 15341
379 W V N N Y L L G N E P P V F D I L F W N 398

15342 AAC GAC ACC ACG CGC CTG CCG GCC GCC TTC CAC GGC GAC CTG ATC GAA ATG TTC AAG AGC 15401
399 N D T T R L P A A F H G D L I E M F K S 418

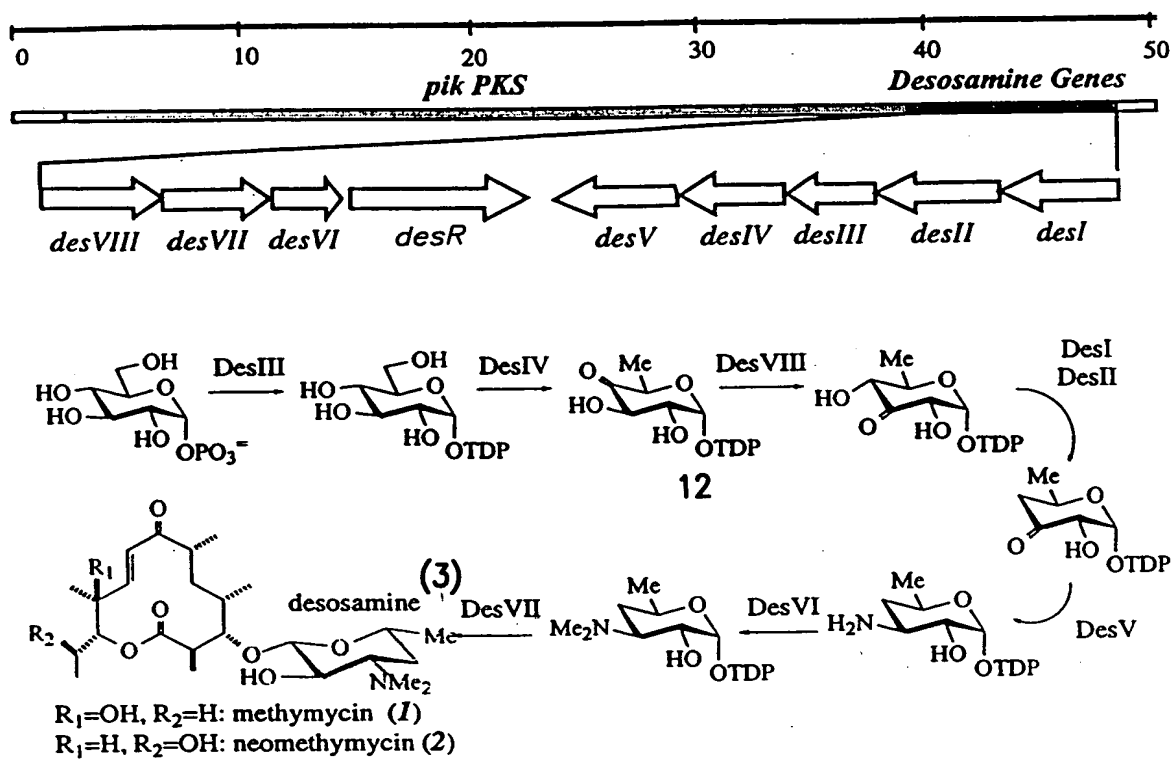
15402 AAC CCG CTG ACC CGC CCG GAC GCC CTG GAG GTT TGC GGC ACT CCG ATC GAC CTG AAA CAG 15461
419 N P L T R P D A L E V C G T P I D L K Q 438

15462 GTC AAA TGC GAC ATC TAC AGC CTT GCC GGC ACC AAC GAC CAC ATC ACC CCG TGG CAG TCA 15521
439 V K C D I Y S L A G T N D H I T P W Q S 458

FIG. 23J

15522	TGC	TAC	CGC	TCG	GCG	CAC	CTG	TTC	GGC	GGC	AAG	ATC	GAG	TTC	GTG	CTG	TCC	AAC	AGC	GGC	15581
459	C	Y	R	S	A	H	L	F	G	G	K	I	E	F	V	L	S	N	S	G	478
15582	CAC	ATC	CAG	AGC	ATC	CTC	AAC	CCG	CCA	GGC	AAC	CCC	AAG	GCG	CGC	TTC	ATG	ACC	GGT	GCC	15641
479	H	I	Q	S	I	L	N	P	P	G	N	P	K	A	R	F	M	T	G	A	498
15642	GAT	CGC	CCG	GGT	GAC	CCG	GTG	GCC	TGG	CAG	GAA	AAC	GCC	ACC	AAG	CAT	GCC	GAC	TCC	TGG	15701
499	D	R	P	G	D	P	V	A	W	Q	E	N	A	T	K	H	A	D	S	W	518
15702	TGG	CTG	CAC	TGG	CAA	AGC	TGG	CTG	GGC	GAG	CGT	GCC	GGC	GAG	CTG	GAA	AAG	GCG	CCG	ACC	15761
519	W	L	H	W	Q	S	W	L	G	E	R	A	G	E	L	E	K	A	P	T	538
15762	CGC	CTG	GGC	AAC	CGT	GCC	TAT	GCC	GCT	GGC	GAG	GCA	TCC	CCG	GGC	ACC	TAC	GTT	CAC	GAG	15821
539	R	L	G	N	R	A	Y	A	A	G	E	A	S	P	G	T	Y	V	H	E	558
15822	CGT	TGA	GCTGCAGCGCCGTTGGCCACCTGCGGGACGCCACGGTGTGAATTC																	15872	
559	R	*																		560	

FIG. 23K



DesI: 4-Dehydrase	DesII: Putative reductase
DesIII: TDP-glucose synthase	DesIV: TDP-glucose-4,6-dehydratase
DesV: Aminotransferase	DesVI: N-methyltransferase
DesVII: Glycosyltransferase	DesVIII: Putative tautomerase
DesR: Glucosidase	

FIG. 24

Scheme 2

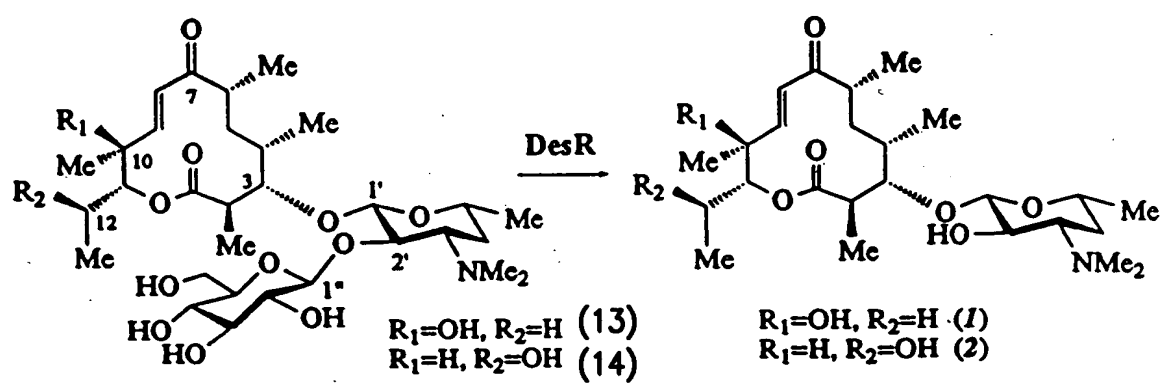


FIG. 25

Scheme 1

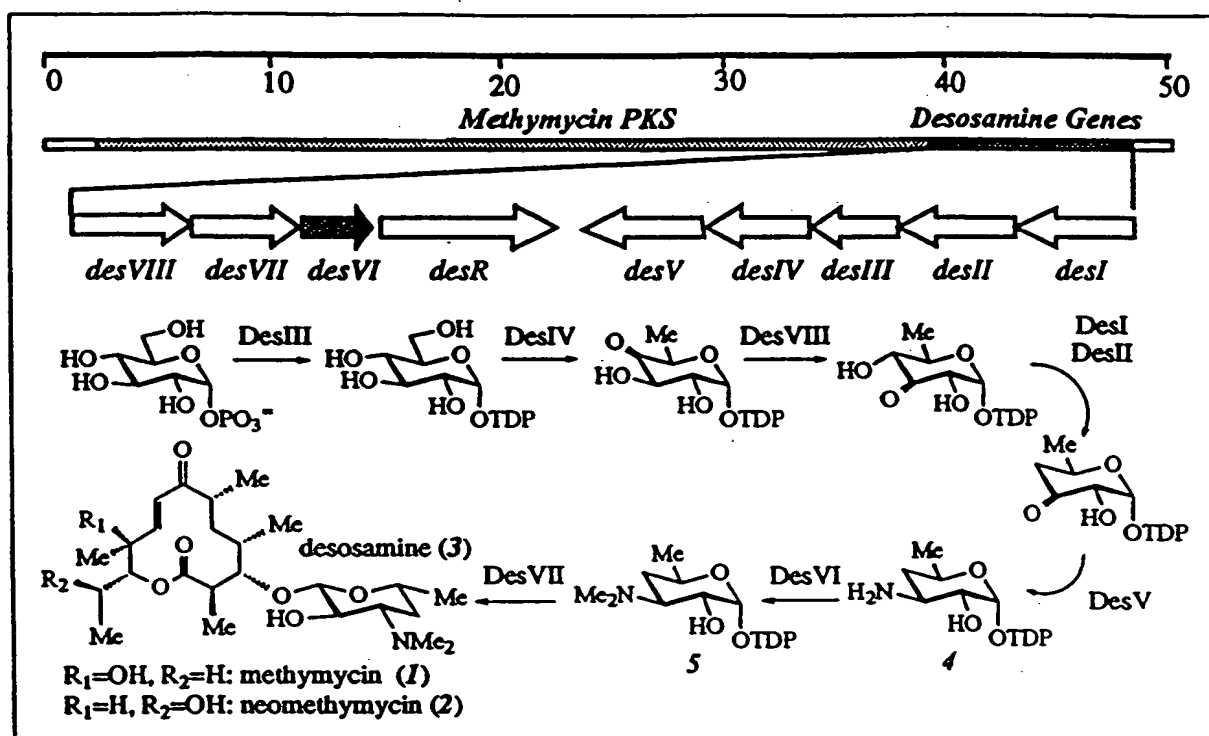


FIG. 26

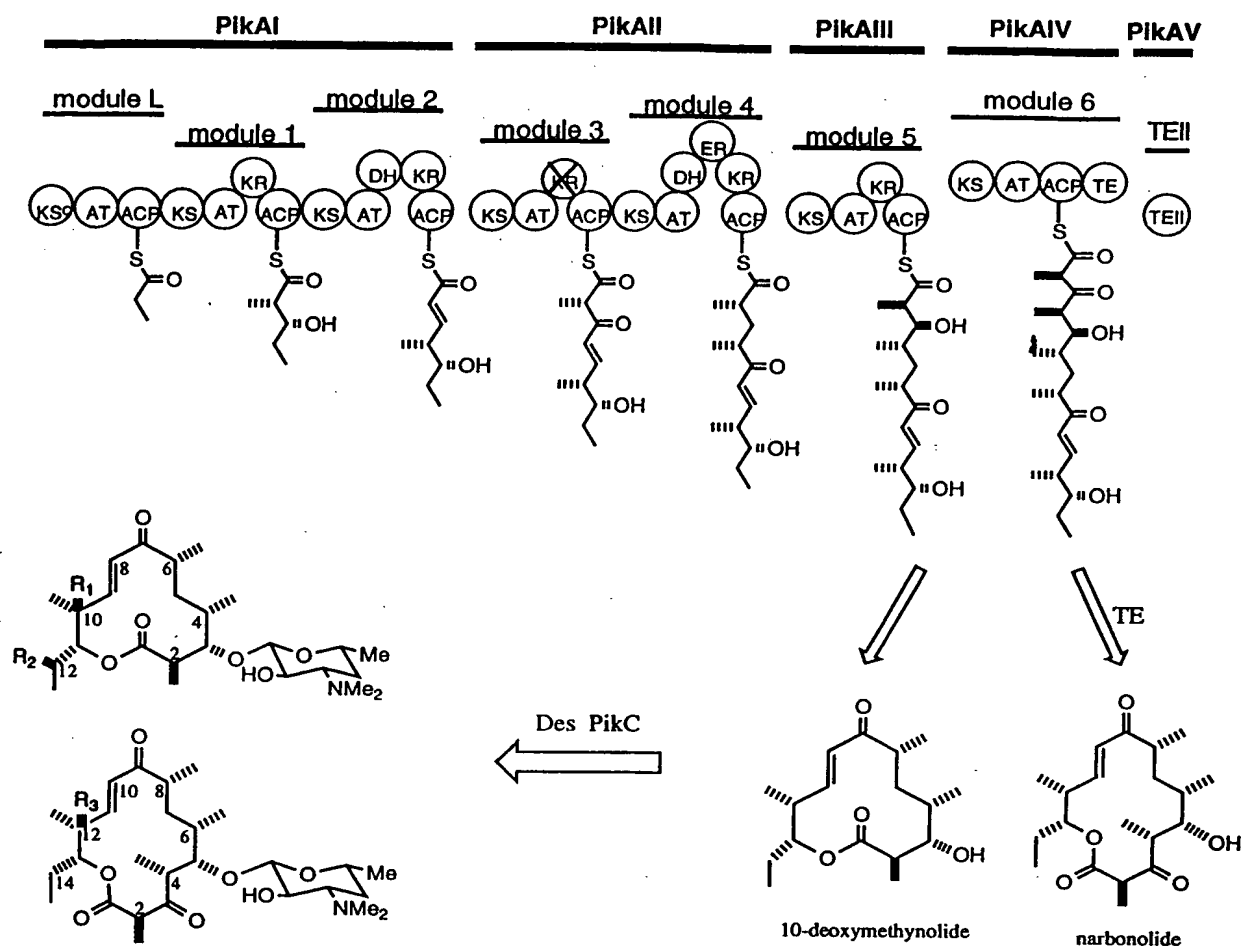


FIG. 28

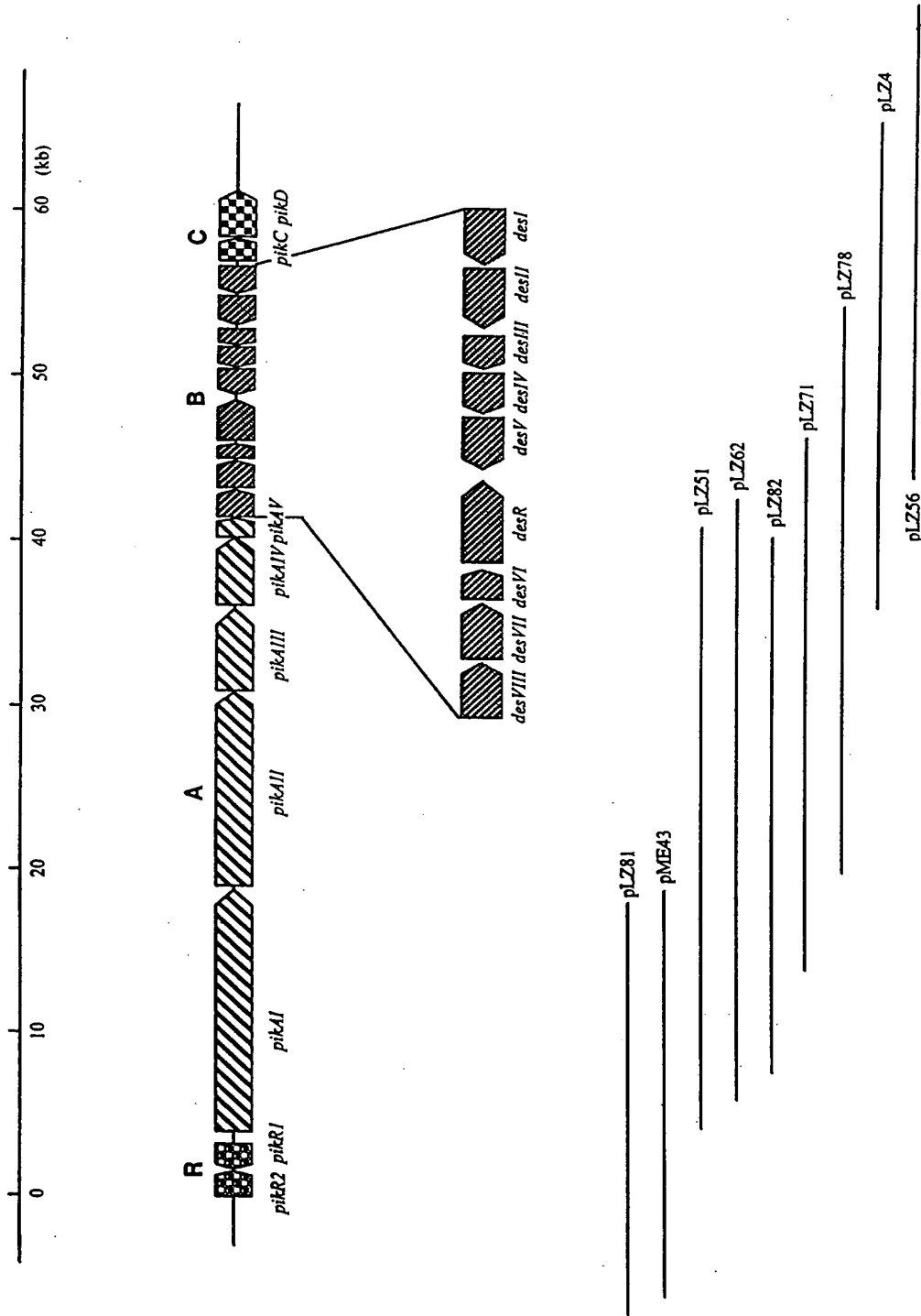


FIG. 29

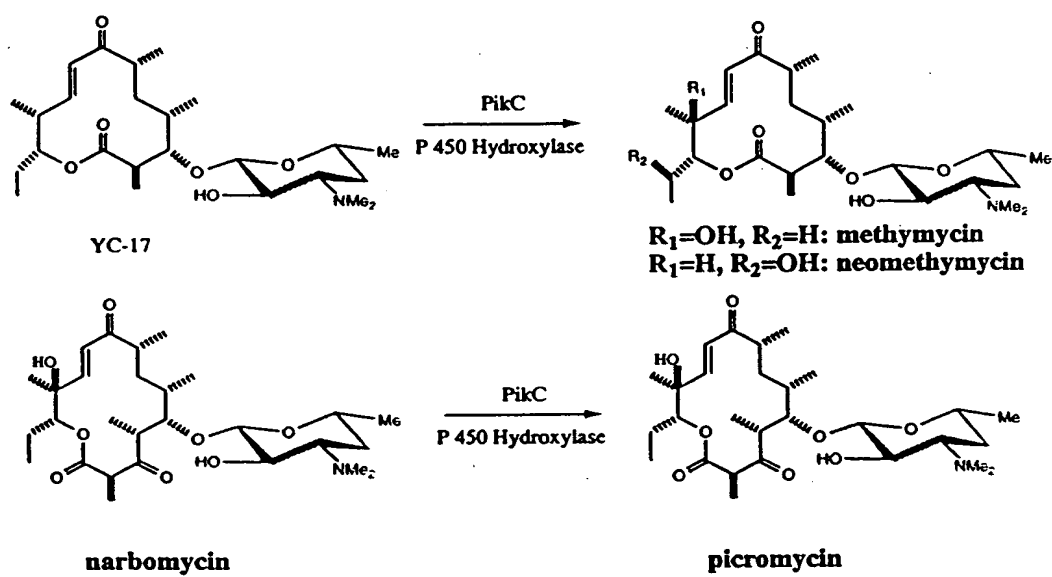


FIG. 30

plkPKS Sequence

10	20	30	40	50	60	70	80	90	100
1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890

BamHI



GGATCCGACC GTGGGTGTGA ATCTCCGGGT GCTCGCCTCG TCCTGCCCGG TTACCTGTCC GCCTCCCGCT CCAGACCAGC GGGAGGCGGA CAGGGGCATG 100

SphI



CCGCCCGGCC GGCTAACGGC CCGTGGGGCG TCCTTACGAC GAGCCTCGCG CGCCTTGGCG GCCCTTGGTC TGCCGGACCT GTGCGCGGGG TCGCCAGGGT 200

BstXI



TGCGCGCGCG GCGTGGGGCC GTATCTGGCG CTCCCGGGCA CGGCGGGCCT GTTCGTCTCC GAGTCATACT CCCTGCCGCC GCGGCGACCG CCCTGGCGCG 300

SphI



GCATGCGCGT GCGCGGGCGC CCGGCGCGCT AACTCGGCTG GGAGGCGCTG AAAAGGGCGA TCCATTGGGT GAGCGTGAGG TCCTTCGGCA GTCCGCGCGTC 400

EcoRI

ApoI



CGGAATTCCG TGGCGGTCCG CGAGGGAACG GTAGGTCCGC TTGGGGATGT GGGCGCGGAG GATCTCCGG AGGCCCGGTC CCGGGCGCGT GAAGACGGCT 500

Figure 31 - 1

plkPKS Sequence

10	20	30	40	50	60	70	80	90	100
1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890
TCGCGCAAGT TCTGGAAGC GCGGCTCGG CTCTCGGCA GCAGGGCTG GGGGCTGCG CTGATCGTCA GGACGCGGCC GTCCACGCGG GGCATCGGAC 600									
GGACGACGA GCGCGGACG CGGTCTGGA CCGCGAATC GTACCAGGG GCCCAGGAG TCGTGAGGAG CGATCCGCGG CTGCGACCGG CGCGTTTGGC 700									
GGCGACTCC CACTGCACTA TCAGGGCGCG CGACTGCCAG TTCTCGATT CCAGGAGACT CCGGAGATC TGGGTCTGTA TGCCGAAGG AACGTTTCCG 800									
ACGACGCTGT CGATATCGG CGGAATGCG AAGTCGAGGA AATCACCCTG GAATACGCTG ACCCTCTCCC CTTCGAATTT CCGCGCGACA TGCGCGGCCC 900									
AGTGGGGTC CATCTCCACG ACGTCAAGG TGTCGAAGGA GCGCACCAAC TCCTCGTTTA TCGCGGCTTT TCGGGGCGG ATTTCGAGAA CGTTCCTACC 1000									

EcoRV BsmI AhoI
 ▼ ▼ BstBI
 ▼ ▼

Figure 31 - 2

plkPKS Sequence

10	20	30	40	50	60	70	80	90	100
1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890

GTCCGCTCG ACATGCGTGA CGAGATTGCG CACGGCTCTG TCGTCCTGAA GGAAGTCTG GCCTAATTG CCGCGAAGG TGTGCGGTC CGCTCGCCTC 1100
 GGTATGGAGT CGCGCATTCG CATgaacgat cccctccctg gatgccgtg tcaatggact tggcacggac catacctcac ggcccgctcg acgaccggag 1200
 XmnI
 aagaagtcca cgcacgggcg ttccggagta cgggagttgt gaacggccgc gacgaagtgc gtcgcggctc ggcgggcggg gacgagcgag gtccggaggga 1300
 acgcgacgaa gcagccgaac cccaagtgc gtgcgacgga gtgacattgg gggcatatcg agggttgtcg tacggagcgc actcaacgag gctccaggag 1400
 ggaggggttg aaccgcgcgc cgactggcct tcgccgcccg cgcggccgga gtatgtcatg tcgggggtga aatcaagcca ttcccccgag atcggctgtt 1500

Figure 31 - 3

pkPKS Sequence

10	20	30	40	50	60	70	80	90	100
1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890

accatccct ttacctggcg tggatttccc aaccttgggt atagagcggg agacgacgcg acaccatgga gaccacgcac accacgagcg ccaccccccg 1600

gccatccccg caaggggggt ccggctcgcc tccgacacc catggcctgg ggtacacgcc aggtataggg ggaacgtagg gggagcatag ggggggtgcc 1700

ctgggggttg gtgaaagcgc ggcttccgga gacggagcg gatgtcttca gccggaatta ccaggaaccg tgcgagaaca ccggtgacag gggtggggc 1800

M S S A G I T R T G A R T P V T G R G A

ggcagcgtgg gacacggggg aagtgcgggt ccgacggggg ttgccccctg ccggccccga tcatggcgag cactccttct ctctgtctcc taccggtgat 1900

A A W D T G E V R V R R G L P P A G P D H A E H S F S R A P T G D

XmnI

EcoRI

ApoI

gtgcccgcg aattgattcg tgagagatg tgcacagtgt ccaagagtga gtccgaggaa ttctgtgtccg tgtcgaacga cgccggttcc gcgcacggca 2000

V R A E L I R G E M S T V S K S E S E E F V S V S N D A G S A H G T

Figure 31 - 4

plkPKS Sequence

10	20	30	40	50	60	70	80	90	100
1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890
CACGGGAACC	CGTCGCCGTC	GTCCGGCATCT	CCTGCCCGGT	GCCCGGCGCC	CGGGACCCGA	GAGAGTTCTG	GGAACCTCCTG	GCGGCAGGCG	GCCAGGCCGT
A E P V A V	V A V V G I S	C R V C R V	P G A R D P R	E F W E L L	A A G G Q A	V			
2100									
CACCGACGTC	CCCGCGGACC	GCTGGAAACG	CGGCGACTTC	TACGACCGG	ACCGCTCCG	CCCGGCGCG	TCGAACAGC	GTTGGGCGG	GTTTCATCGAG
T D V P A D R	W N A G D F	Y D P D R S A	P G R S N S R	W G G F I E					
2200									
GACGTCGACC	GTTTCGACGC	CGCCTTCTTC	GGCATCTCGC	CCCGCGAGGC	CGGGAGATG	GACCCGCGAGC	AGCGGCTCGC	CCTGGAGCTG	GGCTGGGAGG
D V D R F D A	A F F G I S P	R E A A E M	D P Q Q R L A	L E L G W E A					
2300									
CCCTGGAGCG	CGCCGGGATC	GACCCGTCCT	CGCTCACCGG	CACCCGACCG	GGCGTCTTCG	CGGGCGCCAT	CTGGGACGAC	TACGCCACCC	TGAAGCACCG
L E R A G I	D P S S L T G	T R T G V F A	G A I W D D	Y A T L K H R					
2400									
CCAGGGCGGC	GCCGCGATCA	CCCCGCACAC	CGTCACCGGC	CTCCACCGCG	GCATCATCGC	GAACCGACTC	TCGTACACGC	TCGGGCTCCG	CGGCCCCAGC
Q G G A A I	T P H T V T G	L H R G I I A	N R L S Y T L	G L R G P S					
2500									

Figure 31 - 5

plkPKS Sequence

10	20	30	40	50	60	70	80	90	100
1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890

ATGGTGTGTCG ACTCCGGGCCA GTCCTGTGTCG TCCACCTTCG GTGGAGAGC CTGGGGCGCG GCGAGTCCGA GCTCGCCCTC GCCGGCGGCG 2600
M V V D S G Q S S S L V A V H L A C E S L R R G E S E L A L A G G V

TCTCGCTCAA CCTGGTGCG GACAGCATCA TCGGGGCGAG CAAGTTCGCG GGCCTCTCCC CCGACGGCGG TTCGACGCG GCGCCAACGG 2700
S L N L V P D S I I G A S K F G G L S P D G R A Y T F D A R A N G

SnaBI



CTACGTACGC GCGAGGGCG GCGGTTTCGT CGTCTGAAG CGCCTCTCC GGGCCGTGCG CCGGTGCTCG CCGTGATCCG GGGCAGCGCC 2800
Y V R G E G G G F V V L K R L S R A V A D G D P V L A V I R G S A

GTCAACAACG GCGGCGCGC CCAGGGCATG ACGACCCCG GCAGGAGGCC GTGCTCCGCG AGGCCACGA GCGGGCCGG ACCGGCCCG 2900
V N N G G A A Q G M T T P D A Q A Q E A V L R E A H E R A G T A P A

CCGACGTGCG GTACGTCGAG CTGCACGGCA CCGGCACCCC CGTGGGCGAC CCGATCGAG CCGCTGCGCT CCGCGCCCGC CTCGGCACCG GCCGCCCGG 3000
D V R Y V E L H G T G T P V G D P I E A A A L G A A L G T G R P A

Figure 31 - 6

plkPKS Sequence

10	20	30	40	50	60	70	80	90	100
1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890
CGACAGCCG	CTCCTGGTCG	GCTCGGTCAA	GACGAACATC	GGCCACCTGG	AGGGCGCGGC	CGGCATCGCC	GGCCTCATCA	AGGCCGTCCT	GGCGGTCCGC
G Q P	L L V G	S V K T	N I G H	L E G A	A G I A	G L I K	A V L A	V R	
3100									
GGTCGCGCG	TGCCCGCCAG	CCTGAACTAC	GAGACCCCGA	ACCGGCGGAT	CCCGTTCGAG	GAACCTGAAC	TCCGGGTGAA	CACCGAGTAC	CTGCCGTGGG
G R A	L P A S	L N Y E	T P N P	A I P F	E E L N	L R V N	T E Y L	P W E	
3200									
AGCCGGAGCA	CGACGGGCAG	CGGATGGTCG	TCGGCGGTGC	CTCGTTCGGC	ATGGCGGCA	CGAACGCGCA	TGTCGTGCTC	GAAGAGCCC	CCGGGGTTG
P E H	D G Q R	M V V G	V S S F	G M G G	T N A H	V V L E	E A P G	G G C	
3300									
TCGAGGTGCT	TCGGTCGTGG	AGTCGACGGT	CGGCGGGTCG	GCGGTGCGG	GCGGTGTGGT	GCCGTGGGTG	GTGTCGGCGA	AGTCCGCTGC	CGCGCTGGAC
R G A	S V V E	S T V G	G S A V	G G V V	P W V V	S A K S	A A L D		
3400									
GGCAGATCG	AGCGGCTTGC	CGCGTTCGCC	TCGGGGATC	GTACGGATGG	TGTCGACGCG	GGCGCTGTGC	ATCGGGGTGC	TGTCGATGCG	GGTGCTGTCG
A Q I	E R L A	A F A S	R D R T	D G V D	A G A V	D A G A	V D A G	A V A	
3500									

Figure 31 - 7

[illegible]

Figure 31 - 8

pkPKS Sequence

10	20	30	40	50	60	70	80	90	100
1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890

GCCGCCACC	TCGCCGGCAA	GGCGGCATG	CTGTCCCTCG	CGCTGAGCGA	GGACGCCGTC	CTGGAGCGAC	TGGCCGGGTT	CGACGGGCTG	TCCGTGCGCG	4100
A A H L A G K G G M L S L A L S E D A V L E R L A G F D G L S V A A										

CTGTGAACGG	GCCACCGCC	ACCGTGGTCT	CCGGTGACCC	CGTACAGATC	GAAGAGCTTG	CTCGGGCGTG	TGAGGCCGAT	GGGTCCGTG	CGCGGGTCAT	4200
V N G P T A T V V S G D P V Q I E E L A R A C E A D G V R A R V I										

TCCCGTCGAC	TACGCGTCCC	ACAGCCGGCA	GGTCGAGATC	ATCGAGAGCG	AGCTCGCCGA	GGTCTCGCC	GGGCTCAGCC	CGCAGGCTCC	GCGCGTGCCG	4300
P V D Y A S H S R Q V E I I E S E L A E V L A G L S P Q A P R V P										

TTCTTCTCGA	CACTCGAAGG	CGCCTGGATC	ACCGAGCCCG	TGCTCGACGG	CGGCTACTGG	TACCGCAACC	TGCGCCATCG	TGTGGGCTTC	GCCCCGGCCG	4400
F F S T L E G A W I T E P V L D G G Y W Y R N L R H R V G F A P A V										

TCGAGACCT	GGCCACCGAC	GAGGGCTTCA	CCCACTTCGT	CGAGGTGAGC	GCCACCCCG	TCCTCACCAT	GGCCCTCCCC	GGGACCGTCA	CCGGTCTGGC	4500
E T L A T D E G F T H F V E V S A H P V L T M A L P G T V T G L A										

Figure 31 - 9

plkPKS Sequence

10	20	30	40	50	60	70	80	90	100
1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890
GACCTGCGT	CGCGACAACG	GCGGTCAGGA	CGGCTAGTC	GCCTCCCTCG	CCGAAGCATG	GGCCACGGA	CTCGCGGTG	ACTGGAGCCC	GCTCCTCCCC
T L R	R D N G	G G Q D	R L V A	S L A E	A W A	N G L	A V D	W S P	L L P
4600									
TCGCGACCG	GCCACCACTC	CGACCTCCCC	ACCTACGCGT	TCCAGACCGA	GCGCCACTGG	CTGGGCGAGA	TCGAGGCGCT	CGCCCCGGCG	GGCGAGCCGG
S A T G	H H S D	L P T Y	A F Q T	E R H W	L G E I	E A L A	P A G E	P A	
4700									
CGGTGCAGCC	CGCCGTCTC	CGCACGGAGG	CGGCCGAGCC	GGCGGAGCTC	GACCGGACG	AGCAGCTCG	CGTGATCCTG	GACAAGGTCC	GGGCGCAGAC
V Q P	A V L R	T E A A	E P A E	L D R D	E Q L R	V I L D	K V R A	Q T	
4800									
GGCCCAGGTG	CTGGGGTACG	CGACAGGCGG	GCAGATCGAG	GTCGACCGGA	CCTTCCGTGA	GGCCGGTTGC	ACCTCCCTGA	CCGCGGTGGA	CCTGCCGAAC
A Q V	L G Y A	T G G G	Q I E V	D R T F	R E A G	C T S L	T G V D	L R N	
4900									
CGGATCAACG	CGGCCCTTCG	CGTACGGATG	GCGCGGTCCA	TGATCTTCCA	CTTCCCCACC	CCGAGGCTC	TCGGGAGCA	GCTGCTCCTC	GTCGTGCACG
R I N A	A F G V	R M A P	S M I F	D F P T	P E A L	A E Q L	L L L V	V H G	
5000									

MaeI
BfaI

MluI

ApalI

Figure 31 - 10

plkPKS Sequence

10	20	30	40	50	60	70	80	90	100
1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890

GGGAGGCGGC GCGGAACCG GCGGTGCGG AGCCGGCTCC GGTGGCGCG GCGGTGCCG TCGACGAGCC GGTGGCGATC GTCGGCATGG CCTGCCGCCCT 5100
E A A A N P A G A E P A P A A A A G A V D E P V A I V G M A C R L

GCCCCGTGGG GTCGCCTCGC CGGAGGACCT GTGGCGGCTG GTGGCCGGCG GCGGGGACGC GATCTCGGAG TTCCCGCAGG ACCGGCGCTG GGACGTGGAG 5200
P G G V A S P E D L W R L V A G G G D A I S E F P Q D R G W D V E

BamHI
▼

GGGCTGTACC ACCCGGATCC GGAGCACCCC GGCACGTCGT ACGTCCGCCA GGGCGGTTTC ATCGAGAACG TCGCCGGCTT CGACGGCGCC TTCTTCGGGA 5300
G L Y H P D P E H P G T S Y V R Q G G F I E N V A G F D A A F F G I

TCTCGCCCGG CGAGGCCCTC GCCATGGACC CGCAGCAGCG GCTCCTCCTC GAAACCTCCT GGGAGGCCGT CGAGGACGCC GGGATCGACC CGACCTCCCT 5400
S P R E A L A M D P Q Q R L L L E T S W E A V E D A G I D P T S L

GCGGGGACGG CAGGTCGGCG TCTTCACTGG GCGGATGACC CACGAGTACG GGCCGAGCCT GCGGGACGGC GGGGAAGGCC TCGACGGCTA CCTGCTGACC 5500
R G R Q V G V F T G A M T H E Y G P S L R D G G E G L D G Y L L T

Figure 31 - 11

pikPKS Sequence

10	20	30	40	50	60	70	80	90	100
1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890
GGCAACACGG	CCAGCGTGAT	GTCGGGCGCG	GTCTCGTACA	CACCTCGGCT	TGAGGGCCCC	GCCTGACGG	TGGACACGGC	CTGCTCGTCG	TGCTGCTCG
G N T A	S V M S	G R V S	Y T L G	L E G	P A L T	V D T A	C S S S	L V A	
5600									
CCCTGCACCT	CGCCGTGCAG	GCCCTGGCA	AGGGCGAGGT	CGACATGGCG	CTCGCCGGCG	GCGTGGCCGT	GATGCCACAG	CCCGGGATGT	TGCTCGAGTT
L H L A	V Q A L	R K G E	V D M A	L A G G	V A V M	P T P G	M F V E	F	
5700									
CAGCCGGCAG	CGCGGGCTGG	CCGGGGACGG	CCGGTCAAG	GCGTTCCCG	CGTCGGCGGA	CGGCACCAGC	TGGTCCGAGG	GCGTCGGCGT	CCTCCTCGTC
S R Q R	G L A G	D G R S	K A F A	A S A D	G T S W	S E G V	G V L L	V	
5800									
GAGCGCCTGT	CGGACGCCCG	CCGCAACGGA	CACCAGGTCC	TCGCGGTCTGT	CCGCGGCAGC	GCCTGAACC	AGGACGGCGC	GAGCAACGGC	CTCACGGCTC
E R L S	D A R R	N G H Q	V L A V	V R G S	A L N Q	D G A S	N G L T	A P	
5900									
CGAACGGGCC	CTCGCAGCAG	CGCGTCATCC	GGCGGGCGGT	GGCGGACGCC	CGGCTGACGA	CCTCCGACGT	GGACGTCGTC	GAGGCACACG	GCACGGGCAC
N G P S	Q Q R V	I R R A	L A D A	R L T T	S D V D	V V E A	H G T G	T	
6000									

XmnI
▼

Figure 31 - 12

plkPKS Sequence

10	20	30	40	50	60	70	80	90	100
1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890
GGGACTCGGC	GACCCGATCG	AGGCGCAGGC	CCTGATCGCC	ACCTACGGCC	AGGGCCGTGA	CGACGAACAG	CCGCTGCGCC	TCGGGTCGTT	GAAGTCCAAC
R L G	D P I E A Q A	L I A T Y G Q	G R D D E Q	P L R L	G S L	K S N			
6100									
ATCGGGCACA	CCCAGGCGGC	GGCCGGCGTC	TCCGGTGTCA	TCAAGATGGT	CCAGGCGATG	CGCCACGGAC	TGCTGCCGAA	GACGCTGCAC	GTCGACGAGC
I G H T	Q A A A A G V	S G V I K M V	Q A M R H G L	L P K T	L H	V D E	P		
6200									
CCTCGGACCA	GATCGACTGG	TCGGCTGGCG	CCGTGGAACT	CCTCACCGAG	GCCGTGACT	GGCCGGAGAA	GCAGGACGGC	GGGCTGCGCC	GGGCCGCCGT
S D Q	I D W S A G A	V E L L T E	A V D W P E K	Q D G G	L R R	A A V			
6300									
CTCCTCCTTC	GGGATCAGCG	GCACCAATGC	GCATGTGGTG	CTCGAAGAGG	CCCCGGTGGT	TGTCGAGGGT	GCTTCGGTCG	TCGAGCCGTC	GGTTGGCGGG
S S F	G I S G T N A	H V V L E E A	P V V V E G	A S V V	E P S	V G G			
6400									
TCGGCGGTTCG	GCGGCGGTGT	GACGCCTTGG	GTGGTGTTCG	CGAAGTCCGC	TGCCGCGCTC	GACGCGCAGA	TCGAGCGGCT	TGCCGCATTC	GCCTCGCGGG
S A V	G G G V T P W	V V S A K S A	A A L D A Q	I E R L	A A F	A S R	D		
6500									

BsmI



Figure 31 - 13

pkPKS Sequence

10	20	30	40	50	60	70	80	90	100
1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890

ATCGTACGGA TGACGCCGAC GCCGGTGTCTG TCGACGCGGG CGCTGTCTGCT CACGTACTGG CTGACGGGCG TGCTCAGTTC GAGCACCGGG CCGTCGGCGT 6600
R T D D A D A G A V D A G A V A H V L A D G R A Q F E H R A V A L

XmnI



CGGCGCCGGG GCGGACGACC TCGTACAGGC GCTGGCCGAT CCGGACGGGC TGATACGCGG AACGGCTTCC GGTFGCGGCG GAGTGGCGTT CGTGTTCCTCC 6700
G A G A D D L V Q A L A D P D G L I R G T A S G V G R V A F V F P

GGTCAGGGCA CGCAGTGGGC TGGCATGGGT GCCGAACCTGC TGGACTCTTC CGCGGTGTTT CCGGCGGCCA TGGCCGAGTG TGAGGCCGCG CTGTCCCCGT 6800
G Q G T Q W A G M G A E L L D S S A V F A A A M A E C E A A L S P Y

ACGTCGACTG GTCGCTGGAG GCCGTCGTAC GGCAGGCCCC CGGTGGCCCC ACCTGGAGC GGTGCGATGT CGTGCAGCCT GTGACGTTCT CCGTCATGGT 6900
V D W S L E A V V R Q A P G A P T L E R V D V V Q P V T F A V M V

CTCGCTGGCT CGCGTGTGGC AGCACCACGG TGTGACGCCC CAGGCGGTCTG TCGGCCACTC GCAGGGCGAG ATCGCCGCGG CGTACGTCGC CGGAGCCCTG 7000
S L A R V W Q H H G V T P Q A V V G H S Q G E I A A A Y V A G A L

Figure 31 - 14

pkPKS Sequence

10	20	30	40	50	60	70	80	90	100
1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890
CCCCTGGACG	AGCGCGCCCG	CGTCGTCACC	CTGGCGCAGCA	AGTCCATCGC	CGCCGCACTC	GCCGGCAAGG	GCGGCATGCT	GTCCCTCGCG	CTGAACGAGG
P L D D A A R V V T L R S K S I A A H L A G K G G M L S L A L N E D									
ACGCCGTCCT	GGAGCGACTG	AGTGACTTCG	ACGGGCTGTC	CGTCGCGCGC	GTCAACGGGC	CCACCGCCAC	TGTCGTGTCG	GGTGACCCCG	TACAGATCGA
A V L E R L S D F D G L S V A A V N G P T A T V V S G D P V Q I E									
AGAGCTTGCT	CAGGCGTGCA	AGGCGGACGG	ATTCCGCGCG	CGGATCATTC	CCGTGCGACTA	CGCGTCCCAC	AGCCGGCAGG	TCGAGATCAT	CGAGAGCGAG
E L A Q A C K A D G F R A R I I P V D Y A S H S R Q V E I I E S E									
CTCGCCCGAG	TCCTCGCCGG	TCTCAGCCCG	CAGGCCCCCG	GCGTGCCGTT	CTTCTCGACG	CTCGAAGGCA	CCTGGATCAC	CGAGCCCCGTC	CTCGACGGCA
L A Q V L A G L S P Q A P R V P F F S T L E G T W I T E P V L D G T									
CCTACTGGTA	CGCAACCTC	CGTCACCGCG	TGGGTTTCG	CCCGGCCATC	GAGACCCCTGG	CCGTGCGACGA	GGGTTTCACG	CACCTTCGTCG	AGGTCAGCGC
Y W Y R N L R H R V G F A P A I E T L A V D E G F T H F V E V S A									

SphI



MluI



KpnI

Acc65I



Figure 31 - 15

pkPKS Sequence

10	20	30	40	50	60	70	80	90	100
1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890
CGACCCCGTC	CTCACCATGA	CCCTCCCGGA	GACCGTCACC	GGCTCGGCA	CCCTCCGTCG	CGAACAGGGA	GGCCAAAGAGC	GTCTGGTCAC	CTCGCTCGCC
H P V	L T M T	L P E	T V T	G L G T	L R R	E Q G	G Q E R	L V T	S L A
7600									
GAGGCGTGG	TCAACGGGCT	TCCCGTGGCA	TGGACTTCGC	TCTGCGCGC	CACGGCCCTCC	CGCCCGGTC	TGCCACCTA	CGCCTTCCAG	GCCGAGCGCT
E A W V	N G L	P V A	W T S L	L P A	T A S	R P G L	P T Y	A F Q	A E R Y
7700									
ACTGGCTCGA	GAACACTCCC	GCCGCCCTGG	CCACCGGCGA	CGACTGGCGC	TACCGCATCG	ACTGGAAGCG	CCTCCCGGCC	GCCGAGGGGT	CCGAGCGCAC
W L E	N T P	A A L A	T G D	D W R	Y R I D	W K R	L P A	A E G S	E R T
7800									
CGGCCTGTCC	GCGCGCTGGC	TCGCCGTAC	GCCGGAGGAC	CACCTCGCGC	AGCCCGCCGC	CGTGTCTACC	GCGTGTGTCG	ACGCCGGGGC	GAAGGTCTGAG
G L S	G R W L	A V T	P E D	H S A	Q A A	A V L T	A L V D	A G A	K V E
7900									
GTGCTGACGG	CCGGGGCGGA	CGACGACCGT	GAGGCCCTCG	CCGCCCGGCT	CACCGCACTG	ACGACCGGTG	ACGGCTTTCAC	CGGCGTGGTC	TCGCTCCTCG
V L T A	G A D	D D R	E A L A	A R L	T A L	T T G D	G F T	G V V	S L L D
8000									

Eco47III



XhoI
Paer7I
MscI
Bali



Figure 31 - 16

plkPKS Sequence

10	20	30	40	50	60	70	80	90	100
1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890
ACGGACTCGT	ACCGCAGGTC	GCCTGGGTCC	AGGCGCTCGG	CGACGCCGGA	ATCAAGGCGC	CCCTGTGGTC	CGTCACCCAG	GGCGCGGTCT	CCGTGGGACG
G L V	P Q V	A W V	Q A L G	D A G	I K A	P L W S	V T Q	G A V S	V G R
2600	2700	2800	2900	3000	3100	3200	3300	3400	3500
TCTCGACACC	CCCGCCGACC	CCGACCGGGC	CATGCTCTGG	GGCCTCGGCC	GGTGTGTCG	CCTTGAGCAC	CCCGAACGCT	GGGCGGCGCT	CGTCGACCTC
L D T	P A D	P D R	A M L W	G L G	R V V A	L E H	P E R W	A G L	V D L
3600	3700	3800	3900	4000	4100	4200	4300	4400	4500
CCCGCCGAGC	CCGATGCCGC	CGCCCTCGCC	CACCTCGTCA	CCGCACTCTC	CGGGGCCACC	GGCGAGGACC	AGATCGCCAT	CCGCACCACC	GGACTCCACG
P A Q	P D A	A A L A	H L V	T A L S	G A T	G E D	Q I A	I R T	T G L H A
4600	4700	4800	4900	5000	5100	5200	5300	5400	5500
CCCGCGGCCT	CGCCCGGGCA	CCCTCTCCAG	GACGTCGGCC	CACCCGGGAC	TGGCAGCCCC	ACGGCACCGT	CCTCATCACC	GGCGGCACCG	GAGCCCTCGG
R R L	A A R	A P L	H G R R	P T R	D W Q	P H G T	V L I	T G G T	G A L G
5600	5700	5800	5900	6000	6100	6200	6300	6400	6500
CAGCCACGCC	GCACGCTGGA	TGGCCCAACA	CGGAGCCGAA	CACCTCCTCC	TCGTAGCCG	CAGCGGCGAA	CAAGCCCCCG	GAGCCACCCA	ACTCACGCGC
S H A	A A R	W M A	H H G A	E H L L	L L V S	R S G E	Q A P G	A T Q	L T A

BsaBI



Figure 31 - 17

plkPKS Sequence

10	20	30	40	50	60	70	80	90	100
1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890
GAACTCACCG	CATCGGGGCG	CGCGTACAC	ATCGCGGCT	GCGACGTGC	CGACCCCCAC	GCCATGGCA	CCCTCCTCGA	CGCCATCCCC	GCCGAGACGC
E L T A	S G A	R V T	I A A C	D V A	D P H	A M R T	L L D	A I P	A E T P
8600									
CCCTCACCGC	CGTCGTCCAC	ACGCGCGCG	CGCTCGACGA	CGGCATCGTG	GACACGCTGA	CCGCCGAGCA	GGTCCGGCGG	GCCCACCGTG	CGAAGGCCGT
L T A	V V H	T A G	A L D D	G I V	D T L T	A E Q	V R R	A H R	A K A V
8700									
CGGCGCCTCG	GTGCTGACG	AGCTGACCG	GGACCTCGAC	CTCGACGCGT	TCGTGCTCTT	CTCGTCCGTG	TCGAGCACTC	TGGGCATCCC	CGGTCAGGGC
G A S	V L D E	L T R	D L D	L D A F	V L F	S S V	S S T	L G I	P G Q G
8800									
AACTACGCC	CGCACAAACG	CTACCTCGAC	GCCCTCGCGG	CTCGCCGCGG	GGCCACCGGC	CGGTCCGCCG	TCTCGGTGGC	CTGGGGACCG	TGGGACGGTG
N Y A	P H N A	Y L D	A L A	A R R R	A T G	R S A	V S V	A W G	P W D G G
8900									
GCGGCATGGC	CGCCGGTGAC	GGCGTGGCCG	AGCGGCTGCG	CAACCAACGGC	GTGCCCGGCA	TGGACCCGGA	ACTCGCCCTG	GCCGCACCTGG	AGTCCGCGCT
G M A	A G D	G V A	E R L	R N H	G V P	G M D	P E L	A L A	L E S A L
9000									

MluI



Figure 31 - 18

plkPKS Sequence

10	20	30	40	50	60	70	80	90	100
1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890
CGGCCGGAC	GAGACCGCGA	TCACCGTCGC	GGACATCGAC	TGGGACCGCT	TCTACCTCGC	GTA	CTCCTCC	GGTCGCCCCG	AGCCCCCTCGT
G R D	E T A	I T V	A D I	D W D	R F Y L	A Y S	S G R	P Q P	L V E E L
9100									
CGCGAGGTGC	GGCGCATCAT	CGACGCACGG	GACAGCGCCA	CGTCCGGACA	GGGCGGGAGC	TCCGCCCAGG	GCGCCAACCC	CCTGGCCGAG	CGGCTGGCCG
P E V	R I I	D A R	D S A	T S G	Q G G S	S A Q	G A N	P L A	E R L A A
9200									
CGCGGGCTCC	CGGCGAGCGT	ACGGAGATCC	TCCTCGGTCT	CGTACGGGCG	CAGGCCGCCG	CCGTGCTCCG	GATCGGTTCC	CCGGAGGACG	TCGCCGCCGA
A A P	G E R	T E I	L L G	L V R	A Q A	A A V	L R M	R S P	E D V A A D
9300									
CCGCGCCTTC	AAGGACATCG	GCTTCGACTC	GCTCGCCGGT	GTCGAGCTGC	GCAACAGGCT	GACCCGGGCG	ACCGGGCTCC	AGCTGCCCCG	GACGCTCGTC
R A F	K D I	G F D	S L A	G V E	L R N	R L T	R A T	G L Q	L P A T L V
9400									
TTCGACCACC	CGACGCGGCT	GGCCCTCGTG	TCGCTGCTCC	GCAGCGAGTT	CCTCGGTGAC	GAGGAGACGG	CGGACGCCCG	GCGGTCCGCG	GCGCTGCCCG
F D H	P T P	L A L	V S L	L R S	E F L	G D E	E T A	D A R	R S A A L P A
9500									

BstXI
▼

Figure 31 - 19

pkPKS Sequence

10	20	30	40	50	60	70	80	90	100
1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890

BsaBI

CGACTGTGCG TGCCGGTGCC GCGCCGGCG GCGCACCAGC GATCCGATCG CGATCGTCGC GATGAGTGC CGTACCCCG GTGACATCCG 9600
T V G A G A G A G A G T D A D D A D D D P I A I V A M S C R Y P G D I R

CAGCCCGGAG GACCTGTGGC GGATGCTGTC CGAGGGGGCG GAGGGCATCA CGCGTTTCCC CACCGACCGC GGCTGGGACC TCGACGGCCT GTACGACGCC 9700
S P E D L W R M L S E G G E G I T P F P T D R G W D L D G L Y D A

GACCCGGAGC CGCTCGGCAG GCGGTACGTC CCGAGGGCG GGTTCCTGCA CGACGGCGCC GAGTTCGACG CGGAGTTCTT CGGCGTCTCG CCGCGCGAGG 9800
D P D A L G R A Y V R E G G F L H D A A E F D A E F F G V S P R E A

MscI
Bali

CGCTGGCCAT GGACCCGCAG CAGCGGATGC TCCTGACGAC GTCTGGGAG GCCTTCGAGC GGGCCGGCAT CGAGCCGGCA TCGCTGCCG GCAGCAGCAC 9900
L A M D P Q Q R M L L T T S W E A F E R A G I E P A S L R G S S T

CGGTGTCTTC ATCGGCCTCT CCTACCAGGA CTACGGGGCC CGGTCCCGA ACGCCCCCG TGGCGTGGAG GGTACCTGC TGACCGGCAG CACGCCGAGC 10000
G V F I G L S Y Q D Y A A R V P N A P R G V E G Y L L T G S T P S

Figure 31 - 20

pikPKS Sequence

10	20	30	40	50	60	70	80	90	100
1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890

GTCGCGTCGG GCCGTATCGC GTACACCTTC GGTCTCGAAG GGC CGCGGAC GACCGTCGAC ACCGCTGCT CGTCGTCGCT GACCGCCCTG CACCTGGCGG 10100
V A S G R I A Y T F G L E G P A T T V D T A C S S S L T A L H L A V

▼
ApaLI

TGCGGGCGCT GCGCAGCGGC GAGTGCACGA TGGCGCTCGC CGGTGGCGTG GCGATGATGG CGACCCCGCA CATGTCGTG GAGTTCAGCC GTCAGCGGGC 10200
R A L R S G E C T M A L A G G V A M M A T P H M F V E F S R Q R A

GCTCGCCCGG GACGGCCGCA GCAAGGCTT CTCGGCGGAC GCGACGGGT TCGGGCGCGC GGAGGGCGT CCGCTGCTGC TCGTGGAGCG GCTCTCGGAC 10300
L A P D G R S K A F S A D A D G F G A A E G V G L L L V E R L S D

▼
KpnI
Acc65I

AGCGGGCGCA ACGGTCACCC GGTGTCGCC GTGGTCCGCG GTACCGCGGT CAACCAGGAC GCGGCGTAC ACGGCCAAC GGACCTCGC 10400
A R R N G H P V L A V V R G T A V N Q D G A S N G L T A P N G P S Q

AGCAGCGGGT GATCCGGCAG GCGTCGCCG ACGCCCGGCT GGCACCCGCG GACATCGAC CCGTCGAGAC GCACGGCAG GGAACCTCGC TGGCGACCC 10500
Q R V I R Q A L A D A R L A P G D I D A V E T H G T G T S L G D P

Figure 31 - 21

plkPKS Sequence

10	20	30	40	50	60	70	80	90	100
1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890
CATCGAGGCC	CAGGGCCTCC	AGGCCACGTA	CGGCAAGGAG	CGGCCCGCGG	AACGGCCGCT	CGCCATCGGC	TCCGTGAAGT	CCACATCGG	ACACACCCAG
I E A Q G L Q A T Y G K E R P A E R P L A I G S V K S N I G H T Q									
GCCGCGCCG	GTGCGGCGG	CATCATCAAG	ATGGTCCTCG	CGATGCGCCA	CGGCACCCCTG	CCGAAGACCC	TCCACGCCGA	CGAGCCGAGC	CCGCACGTG
A A A G A A G I I K M V L A M R H G T L P K T L H A D E P S P H V D									
ACTGGCGGAA	CAGCGGCCTG	GCCCTCGTCA	CCGAGCCGAT	CGACTGGCCG	GCCGGCACCG	GTCCGGCGCG	CGCCGCCGTC	TCCTCCTTCG	GCATCAGCGG
W A N S G L A L V T E P I D W P A G T G P R R A A V S S F G I S G									
GACGAACGCG	CACGTCGTGC	TGGAGCAGGC	GCCGGATGCT	GCTGGTGAGG	TGCTTGGGGC	CGATGAGGTG	CCTGAGGTGT	CTGAGACGGT	AGCGATGGCT
T N A H V V L E Q A P D A A G E V L G A D E V P E V S E T V A M A									
GGGACGGCTG	GGACCTCCGA	GGTGGCTGAG	GGCTCTGAGG	CCTCCGAGGC	CCCCGGGGCC	CCCGGCAGCC	GTGAGGCGTC	CCTCCCCGGG	CACCTGCCCT
G T A G T S E V A E G S E A S E A P A A P G S R E A S L P G H L P W									

Bsu36I



Figure 31 - 22

pkPKS Sequence

10	20	30	40	50	60	70	80	90	100
1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890

MluI
▼

GGGTGCTGTC CGCCAAGGAC GAGCAGTCGC TCGCGGGCCA GGCCGCCGCC CTGCACGCGT GGCTGTCCGA GCCCGCCGCC GACCTGTGG ACGCCGACGG 11100
V L S A K D E Q S L R G Q A A A L H A W L S E P A A D L S D A D G

ACCGGCCCCG CTGCGGGACG TCGGGTACAC GCTCGGCACG AGCCGTACCG CCTTCGGCGA CCGCGCCGCC GTGACCGCCG CCGACCGGGA CCGGTTCTCG 11200
P A R L R D V G Y T L A T S R T A F A H R A A V T A A D R D G F L

MscI
Bali
▼

GACGGGCTGG CCACGCTGGC CCAGGGCGGC ACCTCGGCC ACCTCCACCT GGACACCGCC CCGGACGGCA CCACCGCGTT CCTCTTCACC GGCCAGGGCA 11300
D G L A T L A Q G G T S A H V H L D T A R D G T T A F L F T G Q G S

EglII
▼

GTACAGCGCC CGGCGCCGCC CGTGAGCTGT ACGACCGGCA CCCGTCTTC GCCGGGCGC TCGACGAGAT CTGCGCCAC CTGACGGTC ACCTCGAACT 11400
Q R P G A G R E L Y D R H P V F A R A L D E I C A H L D G H L E L

GCCCCTGCTC GACGTGATGT TCGCGGCCGA GGGCAGCGCG GAGGCCGCC TGTTCGACGA GACGGGTAC ACGAGTGGC CGTGTTCGC CCTGGAGTC 11500
P L L D V M F A A E G S A E A A L L D E T R Y T Q C A L F A L E V

Figure 31 - 23

pkPKS Sequence

10	20	30	40	50	60	70	80	90	100
1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890

SphI
▼

GGCTCTTCC GGCCTGTCGA GAGCTGGGC ATGGGCGCG CCGCACTGCT CGGTCACTCG GTGGGCGAGA TCGCCGCCG GCACGTCGCC GGTGTGTCT 11600
A L F R L V E S W G M R P A A L L G H S V G E I A A A H V A G V F S

CGCTCGCCGA CGCGGCCCG CTGGTCGCC GTACGCGCG GCTCATGACG GAGCTGCCG CCGGTGGCG GATGCTGCC GTCCAGGCC CGGAGGACGA 11700
L A D A A R L V A A R G R L M Q E L P A G G A M L A V Q A A E D E

GATCCGCGTG TGGCTGAGA CGGAGGAGCG GTACGCGGGA CGTCTGGACG TCGCGCGCGT CAACGGCCCC GAGGCCGCC TCTGTCCG CGACGCGGAC 11800
I R V W L E T E E R Y A G R L D V A A V N G P E A A V L S G D A D

SphI
▼

GCGGCGGGG AGGCGGAGGC GTACTGTCC GGGTCGGCC GCAGGACCG CGCGTCCG GTACGCCAG CCTTCCACTC CGGCACATG GACGGCATG 11900
A A R E A E A Y W S G L G R R T R A L R V S H A F H S A H M D G M L

TCGACGGGT CCGCGCGCTC CTGGAGACGG TGGAGTTCCG GCGCCCTCC CTGACCGTGG TCTCGAAGT CACCGGCCTG GCCGCCGCC CGGACGACCT 12000
D G F R A V L E T V E F R R P S L T V V S N V T G L A A G P D D L

Figure 31 - 24

pkPKS Sequence

10	20	30	40	50	60	70	80	90	100
1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890

ScaI



GTGCGACCCC GAGTACTGGG TCCGGCACGT CCGGGGCACC GTCCGCTTCC TCGACGGCGT CCGTGTCTG CCGGACCTCG GCGTGGGAC CTGCCTGGAG 12100
C D P E Y W V R H V R G T V R F L D G V R V L R D L G V R T C L E

CTGGGCCCCG ACGGGTCCT CACCGCCATG GCGGCCGACG GCCTCGCGGA CACCCCCCGG GATTCGCTG CCGGCTCCCC CGTCGGCTCT CCGGCCGGCT 12200
L G P D G V L T A M A A D G L A D T P A D S A A G S P V G S P A G S

CTCCCGCGGA CTCCGCCGCC GCGGGCTCC GCGGCTCGT GCGTGTCTG GCGCAAGCG GTCGAGACC GAGACCGTCG CCGACGCCCT 12300
P A D S A A G A L R P R P L L V A L L R R K R S E T E T V A D A L

CGGCAGGCG CACGCCACG GCACCGACC CGACTGGAC GCCTGGTTG CCGGCTCCG GCGCACCGC GTGACCTGC CCACGTACTC CTTCCGGCGC 12400
G R A H A H G T G P D W H A W F A G S G A H R V D L P T Y S F R R

GACCGCTACT GGCTGGACG CCGGGCGGC GACACCGCG TGGACACCG CCGCCTCGT CTCGGCACCG CCGACACACC GCTGCTCGC GCGTGGTCA 12500
D R Y W L D A P A A D T A V D T A G L G L G T A D H P L L G A V V S

Figure 31 - 25

pkPKS Sequence

10	20	30	40	50	60	70	80	90	100
1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890

GCCTTCGGGA CCGGACGGC CTGTGCTCA CCGCGCGCT CTCCCTGCGC ACCACCGGT GGCTCGGGA CCACGCGTC CTGGGAGCG TCCTGCTCCC 12600
L P D R D G L L L T G R L S L R T H P W L A D H A V L G S V L L P

CGGCGCGCG ATGGTCGAAC TCGCGGGAG CGCTGGGGG TCGCGCGCT TCGGTGACGT GCGGAGCTG ACCTCCTTG AACCGCTGGT ACTGCCCGAG 12700
G A A M V E L A A H A A E S A G L R D V R E L T L L E P L V L P E

CACGTGGCG TCGAGCTGG CGTGACGGTC GGGCGCGCG CCGGAGGCC CGGTGGCGG TCGGCGGGG ACGGCGCAG GCCCGTCTCC CTCCACTCGC 12800
H G G V E L R V T V G A P A G E P G G E S A G D G A R P V S L H S R

KpnI
▼
Acc65I
▼
MscI
Bali
▼

GGCTCGCGGA CGGCGCGCG GGTACCGCCT GGTCTTGCCA CGGACCGGT CTGTGGCCA CCGACCGGC CGAGTTCCC GTGCGCGCG ACCGTGCGGC 12900
L A D A P A G T A W S C H A T G L L A T D R P E L P V A P D R A A

CATGTGGCG CCGCAGGCG CCGAGGAGT GCGCTCGAC GGTCTCTAC AGCGCTCGA CGGGAACGGC CTCGCCCTCG GTCCGCTGTT CCAGGGGCTG 13000
M W P P Q G A E E V P L D G L Y E R L D G N G L A F G P L F Q G L

Figure 31 - 26

plkPKS Sequence

10	20	30	40	50	60	70	80	90	100
1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890

BsmI



AACGCGGTGT GCGGTACGA GGGTGAGGTC TTCGCCGACA TCGCGCTCCC CGCCACCACG AATGCGACCG CGCCCGCGAC CGCGAACGGC GCGGGGAGTG 13100
N A V W R Y E G E V F A D I A L P A T T N A T A P A T A N G G S A

CGGCGGCGGC CCCCTACGGC ATCCACCCCG CCTTGCTCGA CGCTTCGCTG CACGCCATCG CGGTGCGGG TCTCGTCGAC GAGCCCGAGC TCGTCCGCGT 13200
A A A P Y G I H P A L L D A S L H A I A V G G L V D E P E L V R V

ApalI



CCCCCTCCAC TGGAGGGTG TCACCGTGCA CGCGGCGCGT GCCGGGTCCG TCTCGCTCC CGGGGACGG ACGCCGTCTC GCTGTCCCTG 13300
P F H W S G V T V H A A G A A A R V R L A S A G T D A V S L S L

ACGGACGGCG AGGACGCCC GCTGGTCTCC GTGGAACGGC TCACGCTCG CCCGTCACC GCCGATCAG CGGCGGCGAG CCGGTCGGC GGGCTGATGC 13400
T D G E G R P L V S V E R L T L R P V T A D Q A A A S R V G G L M H

ACGGGTGGC CTGGCGTCG TACGCCCTCG CCTGCTCGG CGAACAGGAC CGCACGCCA CTTCGTACGG GCCGACCGC GTCTCGGCA AGGACGAGCT 13500
R V A W R P Y A L A S S G E Q D P H A T S Y G P T A V L G K D E L

Figure 31 - 27

pkPKS Sequence

10	20	30	40	50	60	70	80	90	100
1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890

GAAGGTGCGC GCGCCCTGG AGTCCGCGGG CGTCGAAGTC GGGCTCTACC CCGACCTGGC CGGCTGTCC CAGGACGTGG CCGCCGGCGC CCGCGGCGCC 13600
K V A A A L E S A G V E V G L Y P D L A A L S Q D V A A G A P A P

CGTACCGTCC TTGCGCGGCT GCCCGCGGGT CCGGCCGACG GCGCGCGGA GGGTGACGG GGACCGTGG CCGGACGCT GGAGCTGCTC CAGGCTGGC 13700
R T V L A P L P A G P A D G G A E G V R G T V A R T L E L L Q A W L

TGGCCGACGA GCACCTCGCG GGCACCGGCC TGCTCCTGGT CACCCGCGGT CCGGTGCGGG ACCCCGAGGG GTCCGGCGCC GACGATGGCG GCGAGGACCT 13800
A D E H L A G T R L L L V T R G A V R D P E G S G A D D G G E D L

NotI
▼

GTCCGACGGC GCGGCTGGG GTCTCGTACG GACCGGCAG ACCGAGAACC CCGCCGCTT CCGCCTTCTC GACCTGGCGG ACGACGCTC GTCGTACCGG 13900
S H A A A W G L V R T A Q T E N P G R F G L L D L A D D A S S Y R

BstXI
▼

ACCTGCGGT CGGTGCTCTC CGACGCGGGC CTGCGGACG AACCGCAGT CGCCCTGCAC GACGGACCA TCAGGCTGGC CCGCTGGCC TCCGTCCGC 14000
T L P S V L S D A G L R D E P Q L A L H D G T I R L A R L A S V R P

Figure 31 - 28

plkPKS Sequence

10	20	30	40	50	60	70	80	90	100
1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890
CCGAGACCGG	CACCGCCGCA	CCGGCGTCTG	CCCCGGAGGG	CACGGTCCTG	CTGACCGGCG	GCACCGGGCG	CCTGGGCGGA	CTGGTCGCCC	GGCACGTGGT 14100
E T G	T A A	P A L A	P E G	T V L	L T G G	T G G	L G G	L V A	R H V V
GGCGAGTGG	GGCGTACGAC	GCCTGCTGCT	GGTGAGCCGG	CGGGGCAAGG	ACGCCCCCGG	CGCGGACGAG	CTCGTGCACG	AGCTGGAGGC	CCTGGGAGCC 14200
G E W	G V R R	L L L	V S R	R G T D	A P G	A D E	L V H E	L E A	L G A
GACGTCCTGG	TGGCGGCGTG	CGACGTCGCC	GACCGCGAAG	CCCTCACCGC	CGTACTCGAC	GCCATCCCGG	CCGAACACCC	GCTCACCGCG	GTCGTCCACA 14300
D V S V	A A C	D V A	D R E A	L T A	V L D	A I P A	E H P	L T A	V V H T
CGGCAGGCGT	CCTCTCCGAC	GGCACCCCTCC	CGTCCATGAC	GACGGAGGAC	GTGGAACACG	TACTGCGGCC	CAAGGTCGAC	GCCGCGTTCC	TCCTCGACGA 14400
A G V	L S D	G T L P	S M T	T E D	V E H V	L R P	K V D	A A F	L L D E
ACTCACCTCG	ACGCCCGCAT	ACGACCTGGC	AGCGTTCGTC	ATGTTCTCCT	CCGCCGCCGC	CGTCTTCGGT	GGCGCGGGGC	AGGGCGCCTA	CGCCGCCGCC 14500
L T S	T P A Y	D L A	A F V	M F S S	A A A	V F G	G A G Q	G A Y	A A A

PmlI



ApalI



Figure 31 - 29

pkPKS Sequence

10	20	30	40	50	60	70	80	90	100
1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890

AACGCCACC TCGACGCCCT CGCCTGGCGC CGCCGGGCGC CCGACTCCC CGCCCTCTCC CTCGGCTGGG GCCTCTGGGC CGAGACCAGC GGCATGACCG 14600
N A T L D A L A W R R R A A G L P A L S L G W G L W A E T S G M T G

GGAGCTCGG CCAGGCGGAC CTGCGCCGGA TGAGCCGCGC GGGCATCGGC GGGATCAGC AGCCGAGGG CATCGGCTC CTCGACGCG CCCTCCGCGA 14700
E L G Q A D L R R M S R A G I G G I S D A E G I A L L D A A L R D

CGACCGCCAC CCGTCTCTGC TGCCCTCTGC GCTCGACGCC GCGGGCTGC GGGACGCGC CGGAAACGAC CCGGCCGAA TCCCGGCGCT CTTCGGGAC 14800
D R H P V L L P L R L D A A G L R D A A G N D P A G I P A L F R D

GTCGTGGCG CCAGGACCGT CCGGGCCCGG CCGTCCGCGG CCTCCGCTC GACGACAGCC GGGACGGCGG GCACGCCGGG GACGGCGGAC GGC CGCGCGG 14900
V V G A R T V R A R P S A A S A S T T A G T A G T P G T A D G A A E

XhoI
PaeR7I

AAACGGCGG GGTACGCTC GCCGACCGG CCGCCACCGT GGACGGGCCC GCACGGCAGC GCCTGCTGCT CGAGTTCGTC GTCGGCGAGG TCGCGAAGT 15000
T A A V T L A D R A A T V D G P A R Q R L L L E F V V G E V A E V

Figure 31 - 30

pkPKS Sequence

10	20	30	40	50	60	70	80	90	100
1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890

ScaI

▼ ACTCGGCCAC GCCCGCGGTC ACCGGATCGA CGCCGAACGG GGCCTTCCTCG ACCTCGGCTT CGACTCCCTG ACCGCCGTG AACTCCGCA CCGGCTCAAC 15100
L G H A R G H R I D A E R G F L D L G F D S L T A V E L R N R L N

TCCGCCGTG GCCTGCCCT CCGGCGACC CTGGTCTTCG ACCACCAAG CCGGCGGCA CTGCTCTCC ACCTGGACG CGAGCTGCCG CGCGGCGCT 15200
S A G G L A L P A T L V F D H P S P A A L A S H L D A E L P R G A S

CGGACCAGGA CGGAGCCGG AACCGGAAC GGAACGAGAA CCGGACGACG GCGTCCCGA GCACGCGCG CTGCTGGCAC AACTGACCCG 15300
D Q D G A G N R N G N E N G T T A S R S T A E T D A L L A Q L T R

CCTGGAAGGC GCCTTGGTGC TGACGGGCTT CTCGGACGCC CCGGGAGCG AAGAAGTCTT GGAGCACCTG CCGTCCCTG GCTCGATGGT CACGGGCGAG 15400
L E G A L V L T G L S D A P G S E E V L E H L R S L R S M V T G E

ACCGGACCG GGACCGCGTC CGGAGCCCG GACGGCGCG GGTCCGGCG CGAGGACCG CCTGGGCGG CCGGGACCG AGCCGGGGG GGGAGTGAGG 15500
T G T G T A S G A P D G A G S G A E D R P W A A G D G A G G G S E D

Figure 31 - 31

plkPKS Sequence

10	20	30	40	50	60	70	80	90	100
1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890

BspHI



ACGGCGCGG AGTGC CGGAC TTCATGAACG CCTCGGCCGA GGAACCTCTTC GGCCTCCTCG ACCAGGACCC CAGCACGGAC TGATCCCTGC CGCACGGTCG 15600
G A G V P D F M N A S A E E L F G L L D Q D P S T D

XmnI



CCTCCCGCCC CGACCCCCGT CCGGGGCACC TCGACTCGAA TCACCTTCATG CGCGCCTCGG GCGCCTCCAG GAACTCAAGG GGAACGCGTG TCCACGGTGA 15700
V S T V N

ACGAAGAGAA GTACCTCGAC TACCTGCGTC GTGCCACGGC GGACCTCCAC GAGGCCCGTG GCCGCTCCG CGAGCTGGAG GCGAAGGCGG GCGAGCCGGT 15800
E E K Y L D Y L R R A T A D L H E A R G R L R E L E A K A G E P V

GGCGATCGTC GGCATGGCCT CCGGCCTGCC CGGCGGCGTC GCCTCGCCCG AGGACCTGTG GCGGCTGGTG GCCGGCGGCG AGGACGCGAT CTCGGAGTTC 15900
A I V G M A C R L P G G V A S P E D L W R L V A G G E D A I S E F

CCCCAGGACC GCGGCTGGGA CGTGGAGGGC CTGTACGACC CGAACCCGGA GGCCACGGGC AAGAGTTACG CCGCGGAGGC CGGATTCTTG TACGAGGCGG 16000
P Q D R G W D V E G L Y D P N P E A T G K S Y A R E A G F L Y E A G

Figure 31 - 32

plkPKS Sequence

10	20	30	40	50	60	70	80	90	100
1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890

CGGAGTTGGA CGCCGACTTC TTGGGATCT CGCGCGCGGA GGCCCTCGCC ATGGACCCGC AGCAGGTCT CCTCCTGGAG GCCTCCTGGG AGCGGTTCGA 16100
E F D A D F F G I S P R E A L A M D P Q Q R L L L E A S W E A F E

BamHI



GCACGCCGGG ATCCCGGCGG CCACCGCGG CGGCACCTCG GTCGGCGTCT TCACCGGCGT GATGTACCAC GACTACGCCA CCCGTCTCAC CGATGTCCCG 16200
H A G I P A A T A R G T S V G V F T G. V M Y H D Y A T R L T D V P

GAGGGCATCG AGGGCTACCT GGGCACCGGC AACTCCGGCA GTGTGCGCTC GGGCCGCGTC GCGTACACGC TTGGCCTGGA GGGGCCGGCC GTCACGGTCG 16300
E G I E G Y L G T G N S G S V A S G R V A Y T L G L E G P A V T V D

ACACCGCCTG CTCGTCTCTG CTGGTCGCCC TGCACCTCGC CGTGCAGGCC CTGCGCAAGG GCGAGGTGGA CATGGCGCTC GCCGGCGGCG TGACGGTTCAT 16400
T A C S S S L V A L H L A V Q A L R K G E V D M A L A G G V T V M

XmnI



GTCGACGCC AGCACCTTCG TCGAGTTCAG CCGTCAGCGC GGGCTGGCGC CGGACGGCCG GTCGAGTCC TTCTCGTCCA CGGCCGACGG CACCAGCTGG 16500
S T P S T F V E F S R Q R G L A P D G R S K S F S S T A D G T S W

Figure 31 - 33

plkPKS Sequence

10	20	30	40	50	60	70	80	90	100
1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890

BamHI



TCCGAGGGCG TCGGGCTCCT CCTCGTCGAG CGCCTGTCCG ACGGCGGTG CAAGGGCCAT CGGATCCTCG CCGTGGTCCG GGCACCGCC GTCAACACAGG 16600
S E G V G V L L V E R L S D A R R K G H R I L A V V R G T A V N Q D

ACGGCGCCAG CAGCGGCCTC ACGGCTCCGA ACGGCGCGTC GCAGCAGCGC GTCATCCGAC GTGCCCTGGC GGACGCCCGG CTCACGACCT CCGACGTGGA 16700
G A S S G L T A P N G P S Q Q R V I R R A L A D A R L T T S D V D

CGTCGTCGAG GCCCACGGCA CCGGTACGCG ACTCGGCGAC CCGATCGAGG CGCAGGCCGT CATCGCCACG TACGGGCAGG GCCGTGACGG CGAACAGCCG 16800
V V E A H G T G T R L G D P I E A Q A V I A T Y G Q G R D G E Q P

BclI



CTGGCCTCG GGTGCTTGAA GTCCAACATC GGACACACCC AGGCCGCCG CCGTGTCTCC GCGGTGATCA AGATGGTCCA GGCATGCGC CACGGCGTCC 16900
L R L G S L K S N I G H T Q A A G V S G V I K M V Q A M R H G V L

PmlI



TGCCGAAGAC GCTCCACGTG GAGAAGCCGA CCGACCAGGT GGA CTGGTCC GCGGCGCGG TCGAGCTGCT CACCGAGGCC ATGGACTGGC CGGACAAGGG 17000
P K T L H V E K P T D Q V D W S A G A V E L L T E A M D W P D K G

Figure 31 - 34

plkPKS Sequence

10	20	30	40	50	60	70	80	90	100
1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890

CGACGGCGGA CTGCGCAGGG CCGCGGTCTC CTCCTTCGGC GTCAGCGGGA CGAACGGCA CGTCGTGCTC GAAGAGGCC CGCGGCCGA GGAGACCCCT 17100
D G G L R R A A V S S F G V S G T N A H V V L E E A P A A E E T P

GCCTCCGAGG CGACCCCGGC CGTCGAGCCG TCGTCCGGC CCGGCCTGGT GCGTGGCTG GTGTGGGA AGACTCCGGC CGGCTGGAC GCCCAGATCG 17200
A S E A T P A V E P S V G A G L V P W L V S A K T P A A L D A Q I G

Asci



GACGCTCGC CGCGTTCGCC TCGCAGGGCC GTACGGACGC CGCCGATCCG GCGCGGTCTG CTCGGTACT GGCCGGCGGG CGCGCCAGT TCGAGCACCG 17300
R L A A F A S Q G R T D A A D P G A V A R V L A G G R A E F E H R

GGCGTCTGT CTCGGCACCG GACAGGACGA TTTCGGCAG GCGTGACCG CTCGGAAGG ACTGATACG GGCACGCCCT CGGACGTGG CCGGTGGCG 17400
A V V L G T G Q D D F A Q A L T A P E G L I R G T P S D V G R V A

XmnI



TTCGTGTTC CCGGTCAGG CACGCAGTGG GCCGGATGG GCGCGAAT CCTCGACGTG TCGAAGGAGT TCGCGCGGC CATGGCCGAG TCGAGAGCG 17500
F V F P G Q G T Q W A G M G A E L L D V S K E F A A A M A E C E S A

Figure 31 - 35

plkPKS Sequence

10	20	30	40	50	60	70	80	90	100
1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890

Asci



CGCTCTCCCG CTAATGTCGAC TGGTCGCTGG AGGCCGTCGT CCGGCAGGCG CCGGGCGCGC CCACGCTGGA GCGGGTCGAC GTGCTCCAGC CCGTGACCTT 17600
L S R Y V D W S L E A V V R Q A P G A P T L E R V D V V Q P V T F

CGTGTCTATG GTTTCGCTGG CGAAGGTCTG GCAGCACCAC GCGGTGACGC CCGAGGCCGT CGTCGGCCAC TCGCAGGGCG AGATCGCGC CGCGTACGTC 17700
A V M V S L A K V W Q H H G V T P Q A V V G H S Q G E I A A A Y V

GCCGTGCCCC TCACCTCGA CGACGCCGCC CGCGTCGTCA CCCTGCGCAG CAAGTCCATC GCCGCCACC TCGCCGGCAA GGGCGGCATG ATCTCCCTCG 17800
A G A L T L D D A A R V V T L R S K S I A A H L A G K G G M I S L A

CCCTCAGCGA GGAAGCCACC CCGCAGCGCA TCGAGAACCT CCACGGACTG TCGATCGCG CCGTCAACGG CCCCACCGCC ACCGTGGTTT CGGGCGACCC 17900
L S E E A T R Q R I E N L H G L S I A A V N G P T A T V V S G D P

CACCCAGATC CAAGAGCTCG CTCAGGCGTG TGAGGCCGAC GGGTCCGCG CACGGATCAT CCCCCTCGAC TAGCCTCCC ACAGCGCCA CGTCGAGACC 18000
T Q I Q E L A Q A C E A D G V R A R I I P V D Y A S H S A H V E T

Figure 31 - 36

plkPKS Sequence

10	20	30	40	50	60	70	80	90	100
1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890

EspI

Bpu1102I

Bsu36I

ATCGAGAGCG AACTCGCCGA GGTCCTCGCC GGGCTCAGCC CGCGGACACC TGAGGTGCCG TTCTTCTCGA CACTCGAAGG CGCCTGGATC ACCGAGCCGG 18100
I E S E L A E V L A G L S P R T P E V P F F S T L E G A W I T E P V

KpnI

Acc65I

TGCTCGACGG CACCTACTGG TACCGCAACC TCCGCCACCG CGTCGGGTTTC GCCCCCGCCG TCGAGACCCCT CGCCACCGAC GAAGGCTTCA CCCACTTCAT 18200
L D G T Y W Y R N L R H R V G F A P A V E T L A T D E G F T H F I

CGAGGTCAGC GCCCACCCTG TCCTCACCAT GACCCTCCCC GAGACCCTCA CCGGCCTCGG CACCCTCCGC CGCGAACAGG GAGGCCAGGA GCGTCTGGTC 18300
E V S A H P V L T M T L P E T V T G L G T L R R E Q G G Q E R L V

ACCTCACTCG CCGAAGCCTG GACCAACGSC CTCACCATCG ACTGGGGCC CGTCCTCCCC ACCGCAACCG GCCACCACCC CGAGCTCCCC ACCTAGCCT 18400
T S L A E A W T N G L T I D W A P V L P T A T G H H P E L P T Y A F

TCCAGCGCCG TCACTACTGG CTCCACGACT CCCC CGCCGT CCAGGCTCC GTGCAGGACT CCTGGCGCTA CCGCATCGAC TGAAGCGCC TCGCGGTGCG 18500
Q R R H Y W L H D S P A V Q G S V Q D S W R Y R I D W K R L A V A

Figure 31 - 37

plkPKS Sequence

10	20	30	40	50	60	70	80	90	100
1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890

MluI
▼

CGACGCGTCC GAGCGGCGCG GGCTGTCCCG GCGGTGGCTC GTCGTCTGCC CCGAGGACCG TTCGGCCGAG GCCGCCCGG TGCTGCCCGC GCTGTCCGCG 18600
D A S E R A G L S G R W L V V V P E D R S A E A A P V L A A L S G

GCCGGGCGCG ACCCGGTACA GCTGGACGTG TCCCGGCTGG GCGACCGGCA GCGGCTCGCC GCGACGCTGG GCGAGGCCCT GCGCGCGGCC GGTGAGCGCG 18700
A G A D P V Q L D V S P L G D R Q R L A A T L G E A L A A A G G A V

TCGACGGCGT CCTCTCGCTG CTCGCGTGGG ACGAGAGCGC GCACCCCGGC CACCCCGCCC CCTTACCCCG GGGCACCGGC GCCACCCCTCA CCCTGGTGCA 18800
D G V L S L L A W D E S A H P G H P A P F T R G T G A T L T L V Q

GGCGCTGGAG GACGCGGCGG TCGCGCGCCC GCTGTGGTGC GTGACCCACG GCGCGGTGTC CGTCGGCCCG GCCGACCACG TCACCTCCCC CGCCCAGGCC 18900
A L E D A G V A A P L W C V T H G A V S V G R A D H V T S P A Q A

ATGGTGTGGG GCATGGGCGG GGTCGCGCGC CTGGAGCACC CCGAGCGGTG GGGCGGCCTG ATCGACCTGC CCTCGGACGC CGACCGGGCG GCCCTGGACC 19000
M V W G M G R V A A L E H P E R W G G L I D L P S D A D R A A L D R

Figure 31 - 38

plkPKS Sequence

10	20	30	40	50	60	70	80	90	100
1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890
GCATGACCAC	GGTCCTCGCC	GGCGGTACGG	GTGAGGACCA	GGTCGGCGTA	CGCGCTCCG	GGCTGCTCGC	CGCGCGCCTC	GTCCGGGCCT	CCCTCCCGGC
M T T V L A	G G T G E D Q	V A V R A S	G L L A R R L	V R A S L L A	R R L V R A	S L P A			
19100									
GCACGGCAGC	GCTTCGCCGT	GGTGGCAGGC	CGACGGCAGC	GTGCTCGTCA	CCGGTGCCGA	GGAGCCTGCG	GCCGCCGAGG	CCGCACGCCG	GCTGGCCCCG
H G T A S P	W Q A D G T	V L V T G A	E E P A A A	E A A R R	L A R				
19200									
GACGGGGCCG	GACACCTCCT	CCTCCACACC	ACCCCTCCG	GCAGCGAAGG	CGCCGAAGGC	ACCTCCGGTG	CCGCCGAGGA	CTCCGGCCTC	GCCGGGCTCG
D G A G H L	L L L H T T	P S G S E G	A E G T S G	A A E D S	G L A G L	V			
19300									
TCGCCGAAC	TCGCCGACCTG	GGCGCGACGG	CCACCGTCGT	GACCTGGAC	CTCACGGACG	CGGAGGCGGC	CGCCCGGCTG	CTCGCCGGCG	TCTCCGACGC
A E L A D L	G A T A T V	V T C D L T	D A E A A	A R L L A	G V S D A				
19400									
GCACCGCTC	AGCGCGTCC	TCCACCTGCC	GCCCACCGTC	GACTCCGAGC	CGCTCGCCGC	GACCGACGG	GACGCGGCTCG	CCCGTGTCGT	GACCGCGAAG
H P L S A V	L L H L P P	T V D S E P	L A A T D A	D A L A R	V V T A	K			
19500									

EspI
Bpu1102I

Figure 31 - 39

pikPKS Sequence

10	20	30	40	50	60	70	80	90	100
1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890

GCCACCGCCG CGCTCCACCT GGACCGCCTC CTGCGGGAGG CCGGGGCTGC CGGAGGCCGT CCGCCCGTCC TGGTCTCTTT CTCCTCGGTC GCCGCGATCT 19600
 A T A A L H L D R L L R E A A A A G G R P P V L V L F S S V A A I W

GGGCGGGCGC CGGTCAGGCG GCGTACGCC CTTCTCGAC GGCCTGCGG GTCAGCACCG GGCCGACGGC CCCACCGTGA CCTCGGTGGC 19700
 G G A G Q G A Y A A G T A F L D A L A G Q H R A D G P T V T S V A

CTGAGAGCCC TGGGAGGGCA GCGCGTGCAC CGAGGGTGCG ACCGGGGAGC GGCTGGCCCG CCTCGGCCTG CCCCCCCTCG CGCGCTCACC 19800
 W S P W E G S R V T E G A T G E R L R R L G L R P L A P A T A L T

GCCCTGGACA CCGCGCTCGG CCACGGGCGAC ACCGCCGTCA CGATCGCCGA CGTCGACTGG TCGAGCTTCG CCCCCGGCTT CACCACGGCC CGGCCGGGCA 19900
 A L D T A L G H G D T A V T I A D V D W S S F A P G F T T A R P G T

Asci

CCCTCTCGC CGATCTGCCC GAGGCGGCC GCGCGTCTGA CGAGCAGCAG TCGACGACGG CCGCCGACGA CACGTCCTG AGCCGCGAGC TCGGTGCGCT 20000
 L L A D L P E A R R A L D E Q Q S T T A A D D T V L S R E L G A L

Figure 31 - 40

plkPKS Sequence

10	20	30	40	50	60	70	80	90	100
1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890
CACCGGCGCC	GAACAGCAGC	GCCGATGCA	GGAGTTGGTC	CGGAGCACC	TCGCCGTGGT	CCTCAACCAC	CCCTCCCCCG	AGGCCGTCGA	CACGGGGCGG
T G A E Q Q R R M Q	E L V R E H L A V V	L N H P S P E A V D	T G R						
20100									
GCCTTCGGTG	ACCTCGGATT	CGACTCGCTG	ACGGCGGTG	AGTCCGCAA	CCGCCTCAAG	AACGCCACCG	GCCTGGCCCT	CCCGGCCACT	CTGGTCTTCG
A F R D L G F D S L	T A V E L R N R L K	N A T G L A L P A T	L V F D						
20200									
ACTACCGGAC	CCCCCGGAGC	CTGGCGGAGT	TCCTCCTCGC	GGAGATCCTG	GGCGAGCAGG	CCGGTGCCGG	CGAGCAGCTT	CCGGTGGACG	GCGGGGTGCA
Y P T P R T L A E F	L L A E I L G E Q A G A G	E Q L P V D G G V D							
20300									
CGACGAGCCC	GTGCGGATCG	TCGGCATGGC	GTGCGGCCTG	CCGGGCGGTG	TCGCCTCGCC	GGAGGACCTG	TGGCGGCTGG	TGGCCGGCGG	CGAGGACGCG
D E P V A I V G M A	C R L P G G V A S P	E D L W R L V A G G	E D A						
20400									
ATCTCCGGCT	TCCCGCAGGA	CCGCGGCTGG	GACGTGGAGG	GGCTGTACGA	CCCGGACCCG	GACGCGTCCG	GGCGGACGTA	CTGCCGTGCC	GGTGGCTTCC
I S G F P Q D R G W	D V E G L Y D P D P	D A S G R T Y C R A	G G F L						
20500									

BstXI



MluI



Figure 31 - 41

pkPKS Sequence

10	20	30	40	50	60	70	80	90	100
1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890
TCGACAGGC	GGGCGAGTTC	GACGCCGACT	TCTTCGGGAT	CTCGCCGGCG	GAGGCCCTCG	CCATGGACCC	GCAGCAGCGG	CTCCTCCTCG	AGACCTCCTG
D E A	G E F	D A D F	F G I	S P R	E A L A	M D P	Q Q R	L L L	E T S W
GGAGGCCGTC	GAGGACGCG	GGATCGACCC	GACCTCCCTT	CAGGGGCAGC	AGGTCGGCGT	GTTGCGGGC	ACCAACGGCC	CCCCTACGA	GCCGCTGCTC
E A V	E D A G	I D P	T S L	Q G Q	Q V G V	F A G	T N G P	H Y E	P L L
CGCAACACCG	CCGAGGATCT	TGAGGGTTAC	GTCGGGACGG	GCAACGCCGC	CAGCATCATG	TCGGGCCGTG	TCTCGTACAC	CCTCGGCCTG	GAGGGCCCGG
R N T A	E D L	E G Y	V G T G	N A A	S I M	S G R V	S Y T	L G L	E G P A
								BsmI	
CCGTACCGGT	CGACACCGCC	TGCTCCTCCT	CGCTGGTCGC	CCTGCACCTC	GCCGTGCAGG	CCCTGCGCAA	GGGCGAATGC	GGACTGGCGC	TCGCGGGCGG
V T V	D T A	C S S	S L V A	L H L	A V Q A	L R K	G E C	G L A	L A G G
								XmnI	
TGTGACGGTC	ATGTCGACGC	CCACGACGTT	CGTGGAGTTC	AGCCGGCAGC	GCGGGCTCGC	GGAGGACGGC	CGGTGCAAGG	CGTTGCGCGC	GTCGGCGGAC
V T V	M S T P	T T F	V E F	S R Q	R G L A	E D G	R S K A	F A A	S A D

Figure 31 - 42

pikPKS Sequence

10	20	30	40	50	60	70	80	90	100
1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890

SphI

GGCTTCGGCC CGGCGGAGG CGTCGGCATG CTCCTCGTCG AGCGCTGTC GGACGCCGC CGCAACGGAC ACCGTGTGCT GGCGTCTGTG CGCGGAGCG 21100
G F G P A E G V G M L L V E R L S D A R R N G H R V L A V V R G S A

CGGTCAACCA GGACGGCGG AGCAACGGC TGACCGCCCC GAACGGGCC TCGCAGCAGC GCGTCATCCG GCGCGCGCTC GCGGACGCCC GACTGACGAC 21200
V N Q D G A S N G L T A P N G P S Q Q R V I R R A L A D A R L T T

CGCCGACGTG GACGTCTGTC AGGCCACCG CACGGGCACG CGACTCGGG ACCCGATCGA GGCACAGGCC CTCATCGCCA CCTACGGCCA GGGCGCGGAC 21300
A D V D V V E A H G T G T R L G D P I E A Q A L I A T Y G Q G R D

ACCGAACAGC CGCTGGCCT GGGTCTGTT AAGTCCAACA TCGGACACAC CCAGGCCGCC GCCGTGTCT CCGGCATCAT CAAGATGGTC CAGGCGATGC 21400
T E Q P L R L G S L K S N I G H T Q A A A G V S G I I K M V Q A M R

PmlI

GCCACGGCGT CCTGCCAAG ACGCTCCACG TGGACCGGCC GTCCGACCAG ATCGACTGGT CGGCGGGCAC GGTCGAGCTG CTCACCGAGG CCATGGACTG 21500
H G V L P K T L H V D R P S D Q I D W S A G T V E L L T E A M D W

Figure 31 - 43

pkPKS Sequence

10	20	30	40	50	60	70	80	90	100
1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890

GCCGAGGAAG CAGGAGGGCG GGCTGCGCCG CGCGGCGGTC TCCTCCTTCG GCATCAGCGG CACGAACGG CACATCGTGC TCGAAGAAGC CCCGGTCGAC 21600.
P R K Q E G G L R R A A V S S F G I S G T N A H I V L E E A P V D

GAGGACGCC CGCGGACGA GCCGTCGGTC GCGGTGTGG TGCCGTGGCT CGTGTCCCG AGACTCCGG CCGCGCTGGA CGCCACAGATC GGACGCCCTCG 21700
E D A P A D E P S V G G V V P W L V S A K T P A A L D A Q I G R L A

CCGCGTTTCG CTCGCAGGCG CGTACGGACG CCGCCGATCC GGGCGCGGTC GCTCGCGTAC TGGCCGGCGG GCGTGCGCAG TTCGAGCACC GGGCCGTGCG 21800
A F A S Q G R T D A A D P G A V A R V L A G G R A Q F E H R A V A

NotI Bsu36I

GCTCGGCACC GGACAGGACG ACCTGGCGGC CGCACTGGCC GCGCCTGAGG GTCTGGTCCG GGGTGTGGCC TCCGTTGTGG GTCGAGTGGC GTTCGTGTTTC 21900
L G T G Q D D L A A A L A A P E G L V R G V A S G V G R V A F V F

XmnI

CCGGGACAGG GCACGCAGTG GGCCGGGATG GGTGCCGAAC TCCTCGACGT GTCGAAGGAG TTCGGGGCGG CCATGGCCGA GTGCGAGGCC GCGCTCGCTC 22000
P G Q G T Q W A G M G A E L L D V S K E F A A A M A E C E A A L A P

Figure 31 - 44

plkPKS Sequence

10	20	30	40	50	60	70	80	90	100
1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890

AscI
▼

CGTACGTGGA CTGGTCGCTG GAGGCGGTCG TCCGACAGGC CCCCGGCGCG CCCACGCTGG AGCGGGTCGA TGTCGTCCAG CCCGTGACGT TCGCCGTGAT 22100
Y V D W S L E A V V R Q A P G A P T L E R V D V V Q P V T F A V M

GGTCTCGCTG GCGAAGGTCT GGCAGCACCA CGGGGTGACC TCGTCGGCCA CTCGCAGGGC GAGATCGCCG CCGCGTACGT CGCCGGTGCC 22200
V S L A K V W Q H H G V T P Q A V V G H S Q G E I A A A Y V A G A

SphI
▼
EspI
Bpu1102I
▼

CTGAGCCTGG ACGACGCCG TCGTGTGCTG ACCCTGCGCA GCAAGTCCAT CGGCGCCAC CTCGCGGGCC AGGGCGGCAT GCTGTCCCTC GCGCTGAGCG 22300
L S L D D A A R V V T L R S K S I G A H L A G Q G G M L S L A L S E

AGGCGGCCGT TGTGAGCGA CTGGCCGGGT TCGACGGGCT GTCCGTGCTC GCCGTCAACG GGCCTACCGC CACCGTGGTT TCGGGCGACC CGACCCAGAT 22400
A A V V E R L A G F D G L S V A A V N G P T A T V V S G D P T Q I

CCAAGAGCTC GCTCAGGCGT GTGAGGCGCA CGGGGTCCGC GCACGATCA TCCCGTCCA CTAGCCTCC CACAGCGCCC ACGTCGAGAC CATCGAGAGC 22500
Q E L A Q A C E A D G V R A R I I P V D Y A S H S A H V E T I E S

Figure 31 - 45

plkPKS Sequence

10	20	30	40	50	60	70	80	90	100
1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890

GAACTGCGCG AGTCCTCTGGC GGGGTTGTCC CCCCAGACAC CCCAGGTCCC CTTCCTCTCC ACCCTCGAAG GGCCTGGAT CACCGAACCC GCCCTCGAGC 22600
 E L A D V L A G L S P Q T P Q V P Q F F S T L E G A W I T E P A L D G

KpnI
 Acc65I

MscI
 BalI

GCGGTACTG GTACCGCAAC CTCGCCCATC GTGTGGGCTT CGCCCGGCC GTCGAAACCC TGGCCACCGA CGAAGGCTTC ACCACTTCG TCGAGGTCAG 22700
 G Y W Y R N L R H R V G F A P A V E T L A T D E G F T H F V E V S

CGCCCAACCC GTCCTCACCA TGGCGCTGCC CGAGACCGTC ACCGGACTCG GCACCCCTCCG CCGTGACAAC GCGGACAGC ACCGCCCTCAC CACCTCCCTC 22800
 A H P V L T M A L P A L P E T V T G L G T L R R D N G G Q H R L T T S L

GCGAGGCTT GGGCCAACGG CCTACCGTC GACTGGGCTT CTCTCTCTCC CACCACGACC ACCACCCCG ATCTGCCAC CTACGCCCTC CAGACCGAGC 22900
 A E A W A N G L T V D W A S L L P T T T T H P D L P T Y A F Q T E R

Eco47III

GCTACTGGCC GCAGCCCGAC CTCTCGCGCG CCGGTGACAT CACCTCCGCC GGTCTCGGG GCGCCGAGCA CCGGTGCTC GCGCGGCGG TGGCGCTCGC 23000
 Y W P Q P D L S A A G D I T S A G L G A A E H P L L G A A V A L A

Figure 31 - 46

plkPKS Sequence

10	20	30	40	50	60	70	80	90	100
1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890

GGACTCCGAC GGCTGCCTGC TCACGGGGAG CCTCTCCCTC CGTACGCACC CTTGGCTGGC GGACCACCGG GTGGCCGGCA CCGTCTGTCT GCCGGGAACG 23100
D S D G C L L T G S L S L R T H P W L A D H A V A G T V L L P G T

GCGTTCGTGG AGCTGGCGTT CCGAGCCGGG GACCAGGTGC GTTGGATCT GGTGAGGAG CTCACCTCG ACGCGCGCT CGTGCTGCCC CGTCGTGGCG 23200
A F V E L A F R A G D Q V G C D L V E E L T L D A P L V L P R R G A

GCGTCCGTGT GCAGCTGTCC GTCGGCGCGA GCGACGAGTC CCGGCGTGT ACCTTCGGGC TCTACGCGCA CCGGAGGAC GCGCGGGCG AGGCGGAGTG 23300
V R V Q L S V G A S D E S G R R T F G L Y A H P E D A P G E A E W

GACGCGGCAC GCCACCGGTG TGCTGGCCGC CCGTGCGGAC CGCACC GCCC CCGTCCCGA CCGGAGGCC TGGCCGCCGC CCGGCGCGCA GCCGTGGAC 23400
T R H A T G V L A A R A D R T A P V A D P E A W P P P G A E P V D

Eco47III

GTGACGGTC TGTACGAGC CTTCCGGCG AACGGCTACG GCTACGGCCC CCTCTTCCAG GCGTCCGTG GTGCTGGCG CCGTGGCGAC GAGGTGTTG 23500
V D G L Y E R F A A N G Y G Y G P L F Q G V R G V W R R G D E V F A

Figure 31 - 47

[illegible]

CCGGGGCGTT CGGGCGGGC ACGCGGTGC CGTTCGCCTG GAGCGGGATC TCCTGACGC GGTGCGGCGC ACCGCCCTCC GCGTGGGCTT GGCCCCCGCC 23700
R G V R R G H A A A V R L E R D L L Y A V G A T A L R V R L A P A

GGCCCGGACA CGGTGTCCGT GAGCGCCGCC GACTCTCTCCG GGACGCCCGT GTTCCGCCGG GACTCCCTCA CGTGCTGCC CGTCACCCC GCGAGCTGG 23800
G P D T V S V S A A D S S G Q P V F A A D S L T V L P V D P A Q L A

CGGCCCTTCAG CGACCCGACT CTGGACGCGC TGCACCTGCT GGAGTGGACC GCCTGGACG GTCCCGCGCA GGCCTGCCC GCGCGGTGCG TGCTGGCGCG 23900
A F S D P T L D A L H L L E W T A W D G A A Q A L P G A V V L G G

CGACGCCGAC GGTCGCGG CGGCGCTGG CGCCGGTGGC ACGAGGTCC TGTCCTTCCC GGACCTTACG GACCTGGTGG AGCCGTCGA CCGGGCGAG 24000
D A D G L A A A L R A G G T E V L S F P D L T D L V E A V D R G E

Figure 31 - 48

plkPKS Sequence

10	20	30	40	50	60	70	80	90	100
1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890
ACCCCGGCC	CGGCGACCGT	CCTGGTGGCC	TGCCCGCGCG	CCGGCCCCCGA	TGGGCCGGAG	CATGTCGGG	AGGCCCTGCA	CGGCTCGTC	GGCTGATGC
T P A P	A T V L V A	C P A A	C P A A	G P D G P E	H V R E	A L H G	S L A	L M Q	
AGGCTGGCT	GGCCGACGAG	CGGTTACCG	ATGGGCGCCT	GGTGCTCGTG	ACCCGCGACG	CGGTGCGCG	CCGTTCCGGC	GACGGCCTGC	GGTCCACGGG
A W L A	D E R F T D	G R L V L V	T R D A	V A A R	S G D G	L R S T	G		
ACAGGCGCC	GTCTGGGGC	TCGGCCGGTC	CGGCGACAG	GAGAGCCCG	GCCGGTTCGT	CCTGCTCGAC	CTCGCCGGG	AAGCCCGGAC	GGCCGGGGAC
Q A A V	W G L G R S	A Q T E S P	G R F V	L L D L	A G E A	R T A	G D		
GCCACCGCC	GGGACGGCCT	GACGACCGG	GACGCCACCG	TCGGCGGCAC	CTCTGGAGAC	GCCGCCCTCG	GCAGCGCCCT	CGCGACCGCC	CTCGGCTCGG
A T A G	D G L T T G	D A T V G G	T S G D	A A L G	S A L A	T A L	G S G		
GCGAGCCGA	GCTCGCCCTC	CGGACGGG	CGTCTCTCGT	ACCCCGCCTG	GGCGGGCCG	CCGGGCCCGC	CGGGCCGAC	GGCCTCGCCG	CGGCCGACGG
E P Q L	A L R D G A	L L V P R L	A R A A	A P A A	A D G L	A A A	D G		

Figure 31 - 49

pikPKS Sequence

10	20	30	40	50	60	70	80	90	100
1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890

CCTCGCGCT CTGCCGCTGC CCGCGGCTCC GGCCCTCTGG CGTCTGGAGC CCGGTACGGA CGGCAGCCTG GAGAGCTCA CGGCGGCGCC CGGCGACGCC 24600
L A A L P L P A A P A L W R L E P G T D G S L E S L T A A P G D A

GAGACCTCG CCCCAGAGCC GCTCGGCGCG GGACAGGTCC GCATCGCGAT CCGGGCCACC GGTCTCAACT TCCGGACGT CCTGATCGCC CTCGGCATGT 24700
E T L A P E P L G P G Q V R I A I R A T G L N F R D V L I A L G M Y

ACCCGATCC GCGGCTGATG GGCACCGAGG GAGCCGGCGT GGTACCGCG ACCGGCCCG GGTACAGCA CCTCGCCCC GGCACCGGG TCATGGGCCT 24800
P D P A L M G T E G A G V V T A T G P G V T H L A P G D R V M G L

GCTCTCCGC GCGTACGCC CCGTCGTCGT GCGGACGCG CGGACCGTCG CCGGATGCC CGAGGGTGG ACGTTCGCC AGGCGCCTC CGTGCCGTG 24900
L S G A Y A P V V V A D A R T V A R M P E G W T F A Q G A S V P V

GTGTTCTGA CGGCCGTCTA CGCCCTGGC GACCTGGCGG ACGTCAAGCC CCGCGAGCG CTCCTGGTCC ACTCCGCCG CGGTGGCGTG GGCATGGCCG 25000
V F L T A V Y A L R D L A D V K P G E R L L V H S A A G G V G M A A

Figure 31 - 50

plkPKS Sequence

10	20	30	40	50	60	70	80	90	100
1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890
CGGTGCAGCT	CGCCCCGGCAC	TGGGGCGTGG	AGGTCCACGG	CACGGCGAGT	CACGGGAAGT	GGGACGCCCT	GCGCGCGCTC	GGCTGGACG	ACGGGCACAT
V Q L A R H	W G V E V H G	T A S H G K W	D A L R A L	G L D D A H	I				
CGCCTCCTCC	CGCACCCCTGG	ACTTCGAGTC	CGCGTTCGT	GCGGGCGGG	CATGGACGTC	GTACTGAAT	CGCTCGCCCC	CGAGTTCGTC	25200
A S S R T L	D F E S A F R	A A S G A G	M D V V L N S	L A R E F V					
GACGCTCGC	TGCGCCTGCT	CGGGCCGGGC	GGCCGGTTCC	TGGAGATGG	GAAGACCGAC	GTCCGGGACG	CGGAGCGGGT	CGCCGCGGAC	CACCCCGGTG
D A S L R L	L G P G G R F V	E M G K T D	V R D A E R V	A A D H P G V					
TCGGCTACCG	CGCCTTCGAC	CTGGGCGAGG	CCGGGCCGGA	GCGGATCGGC	GAGATGCTCG	CCGAGGTCAAT	CGCCCTCTTC	GAGGACGGGG	TGCTCCGGCA
G Y R A F D	L G E A G P E	R I G E M L A	E V I A L F	E D G V L R H					
CCTGCCCGTC	ACGACCTGGG	ACGTGCGCCG	GGCCCCGGAC	GCCTTCGGC	ACGTACGCCA	GGCCCCGCCAC	ACGGGCAAGG	TGCTCCTCAC	GATGCCGTCG
L P V T T W	D V R R A R D	A F R H V S Q	A R H T G K V	V L T M P S					

Figure 31 - 51

plkPKS Sequence

10	20	30	40	50	60	70	80	90	100
1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890

PmlI



GGCCTCGACC CGGAGGGTAC GGTCTCTGCTG ACCGGCGGCA CCGGTGCGT GGGGGGCATC GTGGCCGGC ACGTGGTGG CGAGTGGGC GTACGACGCC 25600
G L D P E G T V L L T G G T G A L G G I V A R H V V G E W G V R R L

ApaLI



TGCTGCTCGT GAGCGGCGG GGCACGGACG CCGCGGCGC CGGCGAGTC GTGCACGAGC TGGAGGCCCT GGGAGCCGAC GTCTCGGTGG CCGCGTCCGA 25700
L L V S R R G T D A P G A G E L V H E L E A L G A D V S V A C D

CGTCGCGAC CGGAGGCC TCACCGCGT ACTCGACTCG ATCCCCGCG AACACCCGCT CACCGCGTC GTCCACACGG CAGGCGTCCT CTCCGACGGC 25800
V A D R E A L T A V L D S I P A E H P L T A V V H T A G V L S D G

ACCTCCCT CGATGACAGC GGAGGATGTG GAACAGTAC TCGTCCCA GGTGACGCC GCGTTCTCTC TCGACGAAC CACCTCGACG CCGGGTACG 25900
T L P S M T A E D V E H V L R P K V D A A F L L D E L T S T P G Y D

ACCTGGCAGC GTTCGTCATG TTCTCTCCG CCGCGGCGT CTTGCGTGGC GCGGCGCAGG GCGCCTACG CGCCGCCAAC GCCACCTCG ACGCCCTCGC 26000
L A A F V M F S S A A A V F G G A G Q G A Y A A A N A T L D A L A

Figure 31 - 52

plkPKS Sequence

10	20	30	40	50	60	70	80	90	100
1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890
CTGGCGCCGC	CGGACAGCCG	GACTCCCGC	CTCTCCCTC	GGCTGGGCC	TCTGGCCGA	GACACGCGC	ATGACCGGC	GACTCAGCGA	CACCGACGC 26100
W R R	R T A G	L P A	L S L	G W G L	W A E	T S G	M T G G	L S D	T D R
TCGGGGCTGG	CCGTTCCGG	GGGACGCC	ATGACAGCG	AGCTGACCCT	GTCCCTCCTG	GACGGGCCA	TGGCGCGCGA	CGACCCGGCG	CTCGTCCCGA 26200
S R L	A R S G	A T P	M D S E	L T L	S L L	D A A	M R R D	D P A	L V P I
TCGCCCTGGA	CGTCGGCGC	CTCCGGCCC	AGCAGCGCA	CGGCATGCTG	GGCCGCTGC	TCAGCGGGCT	CACCCGCGA	TGGGGGTG	GCGGCGGCC 26300
A L D	V A A	L R A Q	Q R D	G M L	A P L L	S G L	T R G	S R V G	G A P
GGTCAACCAG	CGCAGGCAG	CGCCGGAGG	CGCGGCGAG	ACCTCGGCGG	ACCTCGGCGG	GCGGTCGCC	GCGATGACAC	CGGACGACCG	GGTCGGGCAC 26400
V N Q	R R A A	A G G	A G E	A D T D	L G G	R L A	A M T P	D D R	V A H
CTGCGGGACC	TCGTCCGTAC	GCACGTGGC	ACGTCCTGG	GACACGGC	CCCAGCCCG	GTGGACCTGG	AGCGGGCCTT	CCGACACACC	GGTTTCGACT 26500
L R D	L V R T	H V A	T V L G	H G T	P S R	V D L E	R A F	R D T	G F D S

EspI

SphI

Bpu1102I

AscI

PmlI

Figure 31 - 53

pikPKS Sequence

10	20	30	40	50	60	70	80	90	100
1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890
CGTCAACCGC	CGTCAACTC	CGCAACCGTC	TCAACGCCG	GACCGGGCTG	CGGCTGCCGG	CCACGCTGGT	CTTCGACCAC	CCCACCCCGG	GGGAGCTCGC
L T A	V E L	R N R L	N A A	T G L	R L P A	T L V	F D H	P T P	G E L A
CGGGCACCTG	CTCGACGAAC	TGCCACGGC	CGCGGGCGGG	TCCTGGGCGG	AAGGCACCGG	GTCCGGGAGAC	ACGGCCTCGG	CGACCGATCG	GCAGACCACG
G H L	L D E L	A T A	A G G	S W A	E G T	G S G	D T A	S A T	D R Q T T
GGGGCCCTCG	CCGAACTCGA	CCGGCTGGAA	GGCGTGCTCG	CCTCCCTCGC	GCCCGCCGCC	GGCGGCCGTC	CGGAGCTCGC	CGCCCGGCTC	AGGGCGCTGG
A A L	A E L	D R L	E G V	L A S	L A P	A A G	G R P	E L A	A R L A L A
BstXI									
CCGGGGCCCT	GGGGGACGAC	GGCGACGACG	CCACCGACCT	GGACGAGGCG	TCCGACGACG	ACCTCTTCTC	CTTCATCGAC	AAGGAGCTGG	GGGACTCCGA
A A L	G D D	G D D	A T D	L D E	A S D	D D L	F S F	I D K	E L G D S D
CTTCTGACCT	GCCCGACACC	ACCGGCACCA	CCGGCACCCAC	CAGCCCCCCT	CACACACGGA	ACACGGAACG	GACAGGCGAG	AA	CGGGAGCC
F									ATGGCGAACA
									M A N N

Figure 31 - 54

plkPKS Sequence

10	20	30	40	50	60	70	80	90	100
1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890

PstI
SfiI

ACGAAGACAA GCTCCGGCGAC TACCTCAAGC GCGTCACCGC CGAGCTGCAG CAGAACACCA GCGTCTGCG CGAGATCGAG GGACGCACGC ACGAGCCGGT 27100
E D K L R D Y L K R V T A E L Q Q N T R R L R E I E G R T H E P V

GGCGATCGTG GGCATGGCCT GCGCCTGCCC GGGCGGTGTC GCCTGCGCCG AGGACCTGTG GCAGCTGGTG GCGGGGACG GGGACGCGAT CTCGGAGTTC 27200
A I V G M A C R L P G G V A S P E D L W Q L V A G D G D A I S E F

MluI

CCGAGAGACC GCGGCTGGGA CGTGGAGGGG CTGTACGACC CCGACCCGGA CCGGTCCGGC AGGACGTACT GCCGTCCGG CGGATTCCTG CACGACGCCG 27300
P Q D R G W D V E G L Y D P D P D A S G R T Y C R S G G F L H D A G

GCGAGTTTGA CGCCGACTTC TTCCGGATCT CGCCGCGCGA GGCCCTCGCC ATGGACCGC AGCAGCGACT GTCCCTCACC ACCGCGTGG AGGCGATCGA 27400
E F D A D F F G I S P R E A L A M D P Q Q R L S L T T A W E A I E

GAGCGCGGGC ATCGACCGGA CGGCCCTGAA GGGCAGCGGC CTCGGCGTCT TCGTCGGGG CTGGCACACC GGCTACACCT CGGGGCAGAC CACCGCCGTG 27500
S A G I D P T A L K G S G L G V F V G G W H T G Y T S G Q T T A V

Figure 31 - 55

plkPKS Sequence

10	20	30	40	50	60	70	80	90	100
1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890

CAGTCGCCCG AGCTGGAGGG CCACCTGGTC AGCGGCGCGG CGCTGGGCTT CCTGTCCGGC CGTATCGCGT ACGTCCTCGG TACGGACGGA CCGGCCCTGA 27600
Q S P E L E G H L V S G A A L G F L S G R I A Y V L G T D G P A L T

CCGTGGACAC GGCCTGCTCG TCCTCGCTGG TCGCCCTGCA CCTCGCCGTG CAGGCCCTCC GCAAGGGCGA GTGCGACATG GCCCTCGCG GTGGTGTAC 27700
V D T A C S S S L V A L H L A V Q A L R K G E C D M A L A G G V T

XmnI



GGTCATGCCC AACGCGGACC TGTTCTGTGA GTTCAGCCGG CAGCGCGGGC TGGCCGCGGA CGGCCGGTCG AAGCGTTTCG CCACCTCGGC GGACGGCTTC 27800
V M P N A D L F V Q F S R Q R G L A A D G R S K A F A T S A D G F

BamHI



GGCCCCGCGG AGGGCGCCGG AGTCTGCTG GTGGAGCGCC TGTCGGACGC CCGCCGCAAC GGACACCGGA TCCTCGCGGT CGTCCGCGGC AGCGGGTCA 27900
G P A E . G A G V L L V E R L S D A R R N G H R I L A V V R G S A V N

ACCAGGACGG CGCCAGCAAC GGCTCAGCG CTCCGCACGG GGCCTCCAG GGCCTCCAG TCCGACGGC CTGGCGGAC GCCCGGCTCG CGCCGGGTGA 28000
Q D G A S N G L T A P H G P S Q Q R V I R R A L A D A R L A P G D

Figure 31 - 56

pkPKS Sequence

10	20	30	40	50	60	70	80	90	100
1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890
CGTGGACGTC	GTCGAGGCGC	ACGGACACGGG	CACGCGGCTC	GGCGACCCGA	TCGAGGCGCA	GGCCCTCATC	GCCACTACG	GCCAGAGAA	GAGCAGCGAA
V D V	V E A H	G T G	T R L	G D P	I E A Q	A L I	A T Y G	Q E K	S S E
28100									
CAGCCGCTGA	GGCTGGGCGC	GTTGAAGTCG	AACATCGGGC	ACACGCAGGC	CGCGGCCCGT	GTCGCAGGTG	TCATCAAGAT	GGTCCAGGCG	ATGCGCCACG
Q P L R	L G A	L K S	N I G H	T Q A	A A G	V A G V	I K M	V Q A	M R H G
28200									
GACTGCTGCC	GAAGACGCTG	CACGTCGACG	AGCCCTCGGA	CCAGATCGAC	TGGTCGGCGG	GCACGGTGGA	ACTCCTCACC	GAGGCCGTG	ACTGGCCGGA
L L P	K T L	H V D E	P S D	Q I D	W S A G	T V E	L L T	E A V D	W P E
28300									
GAAGCAGGAC	GGCGGGCTGC	GCCGCGCGGC	TGTCTCTCC	TTCGGCATCA	GCGGACGAA	GCGGCACGTC	GTCCTGGAGG	AGGCCCCCGC	GGTCGAGGAC
K Q D	G G L R	A A A	V S S	F G I S	G T N	A H V	V L E E	A P A	V E D
28400									
TCCCCGGCCG	TCGAGCCGCC	GGCCGGTGGC	GGTGTTGGTC	CGTGGCCGGT	GTCGCGGAAG	ACTCCGGCCG	CGCTGGACGC	CCAGATCGGG	CAGCTCGCCG
S P A V	E P P	A G G	G V V	P W P V	S A K	T P A	A L D A	Q I G	Q L A A
28500									

Figure 31 - 57

pikPKS Sequence

10	20	30	40	50	60	70	80	90	100
1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890

BamHI
▼

CGTACGCGGA CGGTCTGTACG GACGTGGATC CGGCCGGTGGC CGCCCCGGCC CTGGTCGACA GCCGTACGCC GATGGAGCAC CGCGCGGTGC CGGTCTGGCGA 28600
Y A D G R T D V D P A V A A R A L V D S R T A M E H R A V A V G D

CAGCCGGGAG GCACTGGGG ACGCCCTGG GATGCGGAA GGA CTGGTAC GCGGCACGTC CTCGGACGTG GGCCGGGTGG CGTTCGTCTT CCCC GGCCAG 28700
S R E A L R D A L R M P E G L V R G T S S D V G R V A F V F P G Q

BsmI
▼

GGCAGCAGT GGGCCGGCAT GGGCGCCGAA CTCCTTGACA GTCACCGGA GTTCGCTGCC TCGATGGCG AATGCGAGAC CGCGCTCTCC CGCTACGTGC 28800
G T Q W A G M G A E L L D S S P E F A A S M A E C E T A L S R Y V D

ACTGGTCTCT TGAAGCCGTC GTCCGACAGG AACCCGGGCGC ACCCAGCTC GACCGGTGC ACGTCGTCCA GCCCGTGACC TTCGTGTCA TGGTCTCGCT 28900
W S L E A V V R Q E P G A P T L D R V D V V Q P V T F A V M V S L

ApalI
▼

GGCGAAGGTC TGGCAGCACC ACGGCATCAC CCCCAGGCC GTCGTGGCC ACTCGCAGGG CGAGATCGCC GCCCGGTACG TCGCCGGTGC ACTCACCTC 29000
A K V W Q H H G I T P Q A V V G H S Q G E I A A A Y V A G A L T L

Figure 31 - 58

pkPKS Sequence

10	20	30	40	50	60	70	80	90	100
1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890
GACGACGCCG	CCCGCGTCGT	CACCTGGCG	AGCAAGTCCA	TGCGCGGCCA	CCTCGCGGC	AAGGGGGCA	TGATCTCCCT	CGCCCTCGAC	GAGCGGCCG
D D A A	R V V T	L R S K	S I A A	H L A G	K G G M	I S L A	L D E A	A A V	
TCCTGAAGCG	ACTGAGCGAC	TTGACGGAC	TCTCCGTGCG	CGCCGTCAAC	GGCCCCACCG	CCACCGTCGT	CTCCGGCGAC	CCGACCCAGA	TCGAGGAAC
L K R L	S D F D	G L S V	A A V N	G P T A	T V V S	G D P T	Q I E E	L	
CGCCCGCACC	TGCGAGGCCG	ACGGCGTCCG	TGCGGGGATC	ATCCCGGTG	ACTACGCCTC	CCACAGCCGG	CAGGTCGAGA	TCATCGAGAA	GGAGCTGGCC
A R T C	E A D G	V R A R	I I P V	D Y A S	H S R Q	V E I I	E K E L	A	
GAGGTCCTCG	CCGGACTCGC	CCCGCAGGCT	CCGCACGTGC	CGTTCTTCTC	CACCCCTCGAA	GGCACCTGGA	TCACCGAGCC	GGTGCTCGAC	GGCACCTACT
E V L A	G L A P	Q A P H	V P F F	S T L E	G T W I	T E P V	L D G T	Y W	
KpnI									
▼									
Acc65I									
▼									
GGTACCGCAA	CCTGCGCCAT	CGCGTGGGCT	TCGCCCCCGC	CGTGGAGACC	TTGGCGGTTG	ACGGCTTCAC	CCACTTCATC	GAGGTCAGCG	CCCACCCCGT
Y R N L	R H R V	G F A P	A V E T	L A V D	G F T H	F I E V	S A H P	V	

Figure 31 - 59

pkPKS Sequence

10	20	30	40	50	60	70	80	90	100
1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890

CCTCACCATG ACCCTCCCG AGACCGTCAC CGGCTCGGC ACCCTCCGCC GCGAACAGGG AGGCCAGGAG CGTCTGGTCA CCTCACTCGC CGAAGCCTGG 29600
L T M T L P E T V T G L G T L R R E Q G G Q E R L V T S L A E A W

Eco47III
▼

GCCAACGGCC TCACCATCGA CTGGGCGCCC ATCCTCCCA CCACCAACCGG CCACCAACCGG GAGCTCCCA CCTACGCCCTT CCAGACCGAG CGCTTCTGGC 29700
A N G L T I D W A P I L P T A T G H H P E L P T Y A F Q T E R F W L

PstI

SfcI

▼

TGCAGAGTC CGGCGCCACC AGCGCGCGG ACGACTGGG TTAGCGGTC GAGTGAAGC CGTGACGGC CTCGGGCCAG GCGGACCTGT CCGGCGGTG 29800
Q S S A P T S A A D D W R Y R V E W K P L T A S G Q A D L S G R W

GATCGTCGCC GTCGGGAGCG AGCCAGAAGC CGAGCTGCTG GCGCGCTGA AGCCCGGGG AGCGGAGGTC GACGTACTGG AAGCCGGGGC GGACGACGAC 29900
I V A V G S E P E A E L L G A L K A A G A E V D V L E A G A D D D

CGTGAGGCC TCGCGCGCG GCTCACCACA CTGACGACCG GCGACGGCTT CACCGGCGTG GTCTCGCTCC TCGACGACCT CGTGCCACAG GTGCGCTGGG 30000
R E A L A A R L T A L T T G D G F T G V V S L L D D L V P Q V A W V

Figure 31 - 60

plkPKS Sequence

10	20	30	40	50	60	70	80	90	100
1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890

TGCAGGCACT CGGCGACGCC GGAATCAAGG CGCCCTTG TCTCCGTCGG ACCTCTCGAC ACCCCGCGCG ACCCCGACCG 30100
Q A L G D A G I K A P L W S V T Q G A V S V G R L D T P A D P D R

GGCCATGCTC TGGGGCCTCG GCGGCTCGT CGCCCTTGAG CACCCCGAAC GCTGGGCGCG CCTCGTCGAC CTCCCGCGCC AGCCCGATGC CGCCGCGCTC 30200
A M L W G L G R V V A L E H P E R W A G L V D L P A Q P D A A A L

BsaBI



GCCACCTCG TCACGCACT CTCGGGCGCC ACGGCGAGG ACCAGATCGC CATCCGCACC ACCGGACTCC AGCCCGCGCG CCTCGCGCGC GCACCCCTCC 30300
A H L V T A L S G A T G E D Q I A I R T T G L H A R R L A R A P L H

ACGGACGTCG GCCACCCGC GACTGGCAGC CCCACGGCAG CGTCCTCATC ACCGGCGGCA CCGGAGCCCT CGGCAGCCAC GCGGACGCT GGATGGCCCA 30400
G R R P T R D W Q P H G T V L I T G G T G A L G S H A A R W M A H

CCACGGAGCC GAACACCTCC TCCTCGTCAG CCGCAGCGGC GAACAAGCC CCGGAGCCAC CCAACTCACC GCCGAATCA CCGCATCGG CGCCCGCGTC 30500
H G A E H L L L V S R S G E Q A P G A T Q L T A E L T A S G A R V

Figure 31 - 61

[illegible]

GGCGACCGGG CGGGCATCCG CTGGACGTCA CCGGCCCGGA GGACATCGCC CGCATCTGG GCGCGAGAC GAGCGGCGCC GAGGTCTCTCG ACGACTGCT 30700
A P G G D P L D V T G P E D I A R I L G A K T S G A E V L D D L L

MLUI

MSCI
Bali

CGTACTGGCA GCGTCGGGGC ATCCGTCCGA TGAGCCCCCGA CCGGGCCCTG GACGAATTGG CCAAGGCCCT GAGCCACGAC GAGACTTCG TCGCCGTGGC 31000
Y W Q R R G I R P M S P D R A L D E L A K A L S H D E T F V A V A

Figure 31 - 62

plkPKS Sequence

10	20	30	40	50	60	70	80	90	100
1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890
CGATGTCGAC	TGGAGCGGT	TCGGCCCCGC	GTTCACGGTG	TCCCGTCCCA	GCCTTCTGCT	CGACGGCGTC	CCGAGGCCCC	GGCAGGCGCT	CGCCGCACCC
D V D W E R F A P A F T V S R P S L L L D G V P E A R Q A L A A P									
31100									
GTCGGTGCCC	CGGCTCCCG	CGACGCGCC	GTGGCGCCGA	CGGGCAGTC	GTGGGCGTG	GCCGCGATCA	CGCGCTCCC	CGAGCCGAG	CGCCGGCCG
V G A P A P G D A A V A P T G Q S S A L A A I T A L P E P E R R P A									
31200									
CGTCTCTCAC	CCTCGTCCGT	ACCCACGCG	CGGCCGTACT	CGGCCATTCC	TCCCCCGACC	GGGTGGCCC	CGGCCGTGCC	TTCACCGAG	TCGGCTTCA
L L T L V R T H A A A V L G H S S P D R V A P G R A F T E L G F D									
31300									
CTCGTGACG	GCCGTGCAG	TCCGCAACCA	GCTCTCCACG	GTGGTCGGCA	ACAGGCTCCC	CGCCACCACG	GTCTTCGACC	ACCCGACGCC	CGCCGCATC
S L T A V Q L R N Q L S T V V G N R L P A T T V F D H P T P A A L									
31400									
GCCGGGCACC	TCCACGAGG	GTACTCGCA	CCGGCCGAGC	CGGCCCGAC	GGACTGGAG	GGCGGGTGC	GCCGGGCCCT	GGCCGAATG	CCCCTCGAC
A A H L H E A Y L A P A E P A P T D W E G R V R R A L A E L P L D R									
31500									

Figure 31 - 63

plkPKS Sequence

10	20	30	40	50	60	70	80	90	100
1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890
GGCTGCGGA	CGCGGGGTC	CTCGACACCG	TCCTGGCCT	CACCGGCATC	GAGCCGAGC	CGGTTCCGG	CGTTTCGGAC	GGCGGCGCCG	CCGACCCCTGG
L R D A G V	L D T V L R L	L R L T G I	E P E P	E P E P	E P E P	G S G G	S D G G	A A D P	G G
31600									
TGCGGAGCCG	GAGGCGTCGA	TCGACGACCT	GGACGCCGAG	GCCCTGATCC	GGATGGCTCT	CGGCCCCCGT	AACACCTGAC	CCGACCGCGG	TCCTGCCCCCA
A E P E A S	I D D L D A E	A L I R M A	L G P R	N T					
31700									
CGCGCCGCAC	CCCGCGCATC	CGCGGCACCA	CCCGCCCCCA	CACGCCCACA	ACCCCATCCA	CGAGCGGAAG	ACCACACCCA	GATGACGAGT	TCCAACGAAC
						M T S S N E Q			
31800									
AGTTGGTGGA	CGCTCTGCGC	GCCTCTCTCA	AGGAGAACGA	AGAACTCCGG	AAAGAGAGCC	GTCCGCCGGC	CGACCGTCGG	CAGGAGCCCA	TGGCGATCGT
L V D A L R	A S L K E N E	E L R E L R	K E S R	R R A D	R R Q E	P M A I	V		
31900									
CGGCATGAGC	TGCCGGTTTG	CGGGCGGAAT	CCGTCCTCCC	GAGGACCTCT	GGGACGCCGT	CGCCGCCGGC	AAGACCTGG	TCTCCGAGGT	ACCGGAGGAG
G M S C R F	A G G I R S P	E D L W D A V	A A G K D L	V S E V	P E E				
32000									

KpnI
▼
Acc65I
▼

Figure 31 - 64

pkPKS Sequence

10	20	30	40	50	60	70	80	90	100
1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890
CGCGGCTGGG	ACATCGACTC	CCTCTAGGAC	CCGGTGCCCG	GGCGCAAGGG	CACGACGTAC	GTCCGCAACG	CCGGTTTCCT	CGACGACGCC	GCCGGATTCC
R G W D	I D S	L Y D	P V P	G R K G	T T Y	V R N A	A F L	D D A	A G F D
ACGGGGCCTT	CTTCGGGATC	TGCGCGCGCG	AGGCCCTCGC	CATGGACCCG	CAGCAGCGGC	AGCTCCTCGA	AGCTCCTGG	GAGTCTTCG	AGCGGGCCCG
A A F	F G I	S P R	E A L	A M D	P Q Q	R Q L	L E A	S W E	V F E
CATCGACCCC	CGTTCGGTCC	GCGGCACCGA	CGTCGGCGTG	TACGTGGGCT	GTGGCTACCA	GGACTACGG	CCGGACATCC	GGGTCGCCCC	CGAAGGCACC
I D P	A S V	R G T	D V G	V Y V	G C G	Y Q D	Y A P	D I R	V A P
GGCGGTTACG	TCGTCACCGG	CAACTCCTCC	GCCGTGGCCT	CCGGGCGCAT	CGCGTACTCC	CTCGGCCCTG	AGGACCCGC	CGTGACCGTG	GACACGGCGT
G G Y V	V T G	N S S	A S A	V A S	G R I	A Y S	L G L	E G P	A V T
GCTCCTTTC	GCTCGTCGCC	CTGCACCTCG	CCCTGAAGGG	CCTGCGGAAC	GGCGACTGCT	CGACGGCACT	CGTGGCGGC	GTGGCCGTCC	TCGCGACGCC
S S S	L V A	L H L	A L K	G L R	N G D	C S T	A L V	G G V	A V L
									A T P

Figure 31 - 65

[illegible]

GTGCGCGGTAC TCTCTCTCGA ACGGCTCTCC GACGCGCGGC GCAAGGGCCA CCGGTCCTTG GCCGTCTGTC GCGGCAGGCG CATCAACACG GACGCGCGCA 32700
V A V L L L E R L S D A R R K G H R V L A V V R G S A I N Q D G A S

GCAACGGCCT CACGCTCCG CACGGGCCT CCCAGCAGCA CCGGCTGG CAGGCCTGG CCGACGGCG GCTCAGTCG AGCAGCTGG ACCTCGTGA 32800
N G L T A P H G P S Q Q H L I R Q A L A D A R L T S S D V D V V E

▶

GGGCCACGGC ACGGGACCC GTCTCGGCGA CCCGATCGAG GCGCAGGCGC TGCTCGCCAC GTACGGGAG GGGCGAGCC CGGGCGCTG 32900
G H G T G T R L G D P I E A Q A L L A T Y G Q G R A P G Q P L R L

GGGACGCTGA AGTGGAAACAT CGGGCACACG CAGGCGCGCTT CGGGTGTTCG CCGGTGTTCG AAGATGTTCG AGCGCTGCG CCACGGGGTG CTGCCAAGA 33000
G T L K S N I G H T Q A A S G V A G V I K M V Q A L R H G V L P K T

Figure 31 - 66

pkPKS Sequence

10	20	30	40	50	60	70	80	90	100
1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890

PmlI



CCCTGCACGT GGACGAGCCG ACGGACCAGG TCGACTGGTC GGCCGGTTTCG GTCGAGCTGC TCACCGAGGC CGTGGACTGG CCGGAGCGGC CCGGCCGGCT 33100
L H V D E P T D Q V D W S A G S V E L L T E A V D W P E R P G R L

CCGCCGGCG GCGTCTCCG CGTTCGGCGT GGGCGGGACG AACGGCTGGA TCGTCTTGA GGAGGCCCG GCGGTCGAGG AGTCCCCTGC CGTCGAGCCG 33200
R R A G V S A F G V G G T N A H V V L E E A P A V E E S P A V E P

CCGGCCGGTG CCGGCGTGGT GCCGTGGCG GTGTCCGCGA AGACTCTGGC CGCACTGGAC GCCCAGATCG GGCAGCTCG CGCATAACGC GAAGACCGCA 33300
P A G G V V P W P V S A K T S A A L D A Q I G Q L A A Y A E D R T

BamHI



CGGACGTGGA TCGGCGGTG GCCGCCCGCG CCCTGGTCTGA CAGCCGTACG GCGATGGAGC ACCGCGCGGT CCGGTCGGC GACAGCCGG AGGCACTGCG 33400
D V D P A V A A R A L V D S R T A M E H R A V A V G D S R E A L R

GGACGCCCTG CCGATGCCG AAGGACTGGT ACGGGGCACG GTCACCGATC CCGGCCGGT GCGTTCGTC TTCCCCCGCC AGGCACCGCA GTGGGCCGGC 33500
D A L R M P E G L V R G T V T D P G R V A F V F P G Q G T Q W A G

Figure 31 - 67

pkPKS Sequence

10	20	30	40	50	60	70	80	90	100
1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890
<div> <div>EcoRI ApoI</div> <div>▼</div> </div>									
<div> <div>BsmI</div> <div>▼</div> </div>									
ATGGGCGCG	AACCTCCTCGA	CAGTCAACCC	GAATTGCGCG	CCGCCATGGC	CGAATGCGAG	ACCGCACTCT	CCCCGTACGT	CGACTGGTCT	CTCGAAGCCG 33600
M G A E	L L D	S S P	E F A A	A M A	E C E	T A L S	P Y V	D W S	L E A V
TCGTCCGACA	GGCTCCCGAGC	GCACCGACAC	TCGACCGCGT	CGACGTCGTC	CAGCCCGTCA	CCTTCGCCGT	CATGGTCTCC	CTCGCCAAGG	TCTGGCAGCA 33700
V R Q	A P S	A P T L	D R V	D V V	Q P V T	F A V	M V S	L A K V	W Q H
CCACGGCATC	ACCCCGGAGG	CCGTCAATCGG	CCACTCCCAG	GGCGAGATCG	CCGCCCGGTA	CGTCGCCGGT	GCCCTCACCC	TCGACGACGC	CGCTCGTGTC 33800
H G I	T P E A	V I G	H S Q	G E I A	A A Y	V A G	A L T L	D D A	A R V
GTGACCTTCC	GCAGCAAGTC	CATCGCCGCC	CACCTGCGCG	GCAAGGCGG	CATGATCTCC	CTCGCCCTCA	GCGAGGAAGC	CACCCGGCAG	CGCATCGAGA 33900
V T L R	S K S	I A A	H L A G	K G G	M I S	L A L S	E E A	T R Q	R I E N
ACCTCCACGG	ACTGTGATC	GCCGCGGTCA	ACGGGCCTAC	CGCCACCGTG	GTTTCGGGCG	ACCCACCCA	GATCCAAGAA	CTTGCTCAGG	CGTGTGAGGC 34000
L H G	L S I	A A V	N G P T	A T V	V S G D	P T Q	I Q E	L A Q	A C E A

Figure 31 - 68

plkPKS Sequence

10	20	30	40	50	60	70	80	90	100
1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890
CGACGGCATC	CGCGCAGGGA	TCATCCCCGT	CGACTACGCC	TCCCACAGCG	CCCACGTCGA	GACCATCGAG	AACGAACCTG	CCGACGTCCT	GGCGGGGTTG 34100
D G I R A R I	I P V D Y A	S H S A H V E	T I E N E L A	D V L A G L					
TCCCCTCCAGA	CACCCAGGT	CCCCTTCTTC	TCCACCTCG	AAGGCACCTG	GATCACCGAA	CCCGCCCTCG	ACGGCGGCTA	CTGGTACCGC	AACCTCGGCC 34200
S P Q T P Q V	P F F S T L E	G T W I T E	P A L D G G Y	W Y R N L R H					
ATCGTGTTGG	CTTCGCCCGG	GCCGTGAGGA	CCCTCGCCAC	CGACGAAGGC	TTCACCCACT	TCATCGAGGT	CAGCGCCAC	CCCGTCCTCA	CCATGACCC 34300
R V G F A P A	V E T L A T	D E G F T H F	I E V S A H	P V L T M T L					
CCCCGACAAG	GTACCGGCC	TGGCCACCT	CCGACGGGAG	GACGGGGAC	AGCACCGCCT	CACCACTCC	CTTGCCGAGG	CCTGGGCCAA	CGGCCCTCGCC 34400
P D K V T G L	A T L R R E	D G G Q H R L	T T S L A E A	W A N G L A					
CTCGACTGGG	CCTCCCTCCT	GCCCGCCACG	GGCGCCTCA	GCCCCGCCGT	CCCCGACCTC	CCGACGTACG	CCTTCCAGCA	CCGCTCGTAC	TGGATCAGCC 34500
L D W A S L L	P A T G A L S	P A V P D L	P T Y A F Q H	R S Y W I S P					

KpnI
▼
Acc65I
▼

MscI
Bali
▼

Figure 31 - 69

pkPKS Sequence

10	20	30	40	50	60	70	80	90	100
1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890
CGCGGGTCC	CGCGAGGCG	CCCGGCACA	CGCTTCGG	GCGGAGGCC	GTCGCCGAGA	CGGGGCTCGC	GTGGGGCCCG	GGTGCCGAGG	ACCTCGACGA 34600
A G P	G E A	P A H T	A S G	R E A	V A E T	G L A	W G P	G A E D	L D E
GGAGGGCCG	CGCAGCGCG	TACTCGCGAT	GGTGATGCG	CAGGCGGCGT	CCGTGCTCCG	GTGCGACTCG	CCCGAAGAGG	TCCCCGTCGA	CCGCCCCGCTG 34700
E G R	R S A	V L A M	V M R	Q A A S	V L R	C D S	P E E V	P V D	R P L
CGGGAGATCG	GCTTCGACTC	GCTGACCGCC	GTCGACTTCC	GCAACCGCGT	CAACCGGCTG	ACCGGTCTCC	AGCTGCCGCC	CACCGTCGTG	TTCCAGCACC 34800
R E I	G F D S	L T A	V D F R	N R V	N R L	T G L Q	L P P	T V V	F Q H P
*									ACP 6
CGACGCCCCGT	CGCGCTCGCC	GAGCGCATCA	GCGACGAGCT	GGCCGAGCGG	AACTGGGCCG	TCGCCGAGCC	GTCGGATCAC	GAGCAGGCGG	AGGAGGAGAA 34900
T P V	A L A	E R I S	D E L	A E R	N W A V	A E P	S D H	E Q A E	E E K
									TE domain
GGCCGCCGCT	CGGCGGGGG	CCCGTCCCG	GGCCGACACC	GGCGCCGGCG	CCGGGATGTT	CCGGGCCCTG	TTCCGGCAGG	CCGTGGAGGA	CGACCCGGTAC 35000
A A A	P A G A	R S G	A D T	G A G A	G M F	R A L	F R Q A	V E D	D R Y

Figure 31 - 70

[illegible]

▶

TCGCCTGGA GCGGCGCAC GCGCGCGCG GCGCGCGGAT GAGI VLV DPHY PPGH QEP I EVW SRQ
RLE RAH GA PPA GAGI VLV DPHY PPGH QEP I EVW SRQ
GGAGCAGGCA 35500

Figure 31 - 71

pkPKS Sequence

10	20	30	40	50	60	70	80	90	100
1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890

MscI
BamI



GCTGGGCGAG GGCCTGTTCG CCGGCGAGCT GGAGCCGATG TCCGATCGC GGCTGCTGGC CATGGGCCGG TACGGCGGCT TCCTGCGCG CCGGCGCGG 35600
L G E G L F A G E L E P M S D A R L L A M G R Y A R F L A G P R P

GGCCGCAGCA GCGGCGCCGT GCTTCTGTGTC CGTGCCTCCG AACCGCTGGG CGACTGGCAG GAGGAGCGGG GCGACTGGCG TGCCCACTGG GACCTTCGCG 35700
G R S S A P V L L V R A S E P L G D W Q E E R G D W R A H W D L P H

ACACCGTCGC GGCGTGCCG GCGGACCACT TCACGATGAT GCGGCGGCCG TGCGCGAGGC CGTCCTCTCC TGGCTCGACG CCATCGAGGG 35800
T V A D V P G D H F T M M R D H A P A V A E A V L S W L D A I E G

BamHI



CATCGAGGGG GCGGCGAAGT GACCGACAGA CCTCTGAACG TGGACAGCGG ACTGTGGATC CCGGCGCTTC ACCCGCGGCC GAACAGCGCG GTGCGGCTGG 35900
I E G A G K V T D R P L N V D S G L W I R R F H P A P N S A V R L V
→ TE2

TCTGCCTGCC GCACGCGGC GGCTCCGCCA GCTACTTCTT CCGCTTCTCG GAGGAGCTGC ACCCTCCGT CGAGGCGCTG TCGGTGCAGT ATCCGGGCGG 36000
C L P H A G G S A S Y F F R F S E E L H P S V E A L S V Q Y P G R

Figure 31 - 72

plkPKS Sequence

10	20	30	40	50	60	70	80	90	100
1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890
CCAGGACCGG	CGTGCCGAGC	CGTGTCTGGA	GAGCGTCGAG	GAGCTCGCG	AGCATGTGGT	CGCGGCCACC	GAACCTTGGT	GGCAGGAGGG	CCGGCTGGCC 36100
Q D R R A E P	C L E S V E	E L A E H V V	A A T E P W W	Q E G R L A					
TTCTTCGGGC	ACAGCCTCGG	CGCCTCCGTC	GCCTTCGAGA	CGGCCCGCAT	CCTGGAACAG	CGGCACGGGG	TACGGCCCGA	GGCCTGTAC	GTCTCCGGTC 36200
F F G H S L G	A S V A F E T	A R I L E Q	R H G V R P E	G L Y V S G R					
*									

plkPKS Sequence

10	20	30	40	50	60	70	80	90	100
1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890	1234567890
CGTACTCCGG CCGCCCACTTC TACCTCAACG ACCAGTGGCA CGAGATCTGC AACGACATCT CCGACCACCT GCTCGTCACC CGCGGCGCGC CCGATGCCCG 36600									
Y S G	G H F	Y L N	D Q W	H E I	C N D I	S D H	L L V	T R G	A P D A R
* CGTCGTGCAG CCCCCGACCA GCCTTATCGA AGGAGCGGCG AGAAGATGGC AGAA CCCACG GTGACCGACG ACCTGACGGG GGCCCTCAGC CAGCCCCCGC 36700									
V V Q	P P T	S L I	E G A	A K R	W Q N	P R			
TGGGCCGCAC CGTCCCGCGG GTGGCCGACC GTGAAC TCGG CACCCACCTC CTGGAGACCC GCGGCATCCA CTGGATCC 36778									

BglII

Ascl

Figure 31 - 74

sugar.finalgene b-1 Sequence

10	20	30	40	50	60	70	80	90
1234567890123456789012345678901234567890123456789012345678901234567890								
BamHI								
▼								
GGATCCGGCGCTTCCACCCCGCGCCGGAACAGCGGGTGGCTGGCTGCCCTGCCGACGCCGGCGGCTCCGCCAGCTACTTCTTCCGCT								90
CCTAGGCCGCGAAGGTGGGGCGCGGCTTGTCCGCCACGCCGACAGACGGACGGCGTCCGCCCGAGCGGTGCGATGAAGAAGCGGA								
I R R F H P A P A P N S A V R L V C L P H A G G S A S Y F F R F								
TCTCGGAGGAGCTGACCCCTCCGTCCGTCGAGGCCCTGTCTGGTGCAGTATCCGGGGCCGCCAGGACCGGGCTGCCGAGCCGTGTCTGGAGAGCG								
AGAGCCTCCTCGACGTGGGAGGCGAGCTCCGGGACAGCCACGTCTATAGGCCCGGGCGGTCTTGCCCGCACGGCTCGGCACAGACCTCTCGC								180
S E E L H P S V E A L S V Q Y P G R Q D R R A E P C L E S V								
NspHI								
▼								
TCGAGGAGCTCGCCGAGCATGTGGTCCGGGCCACCGAACCTGGTGGCAGGAGGCCGGCTGGCCCTTCTTCGGGCACAGCCTCGGCGCCT								270
AGCTCCTCGAGCGGCTCGTACACCGCGCGGTGGCTTGGGACCACTGCTCCCGGCCGACCGGAAGAAGCCCGTGTCTCGGAGCCCGCGGA								
E E L A E H V V A A T E P W W Q E G R L A F F G H S L G A S								
Ascl								
▼								
CCGTCGCCCTTCGAGACGGCCCGCATCCTGGAACAGCGGCACGGGGTACGGCCCGAGGGCCTGTACGTCTCCGGTCGGCGGCCCGGTCGC								360
GGCAGCGGAAGCTCTGCCGGGCGTAGGACCTTGTCCCGGTGCCCATGCCGGGCTCCCGGACATGCAGAGGCCAGCCGCGCGGGCAGCG								
V A F E T A R I L E Q R H G V R P E G L Y V S G R R A P S L								

FIG. 32 - 1

[illegible]

```

EspI
Bpu1102I

```

► **Ides**

► BglII

CGTTCTGCTCCGGGGTACTCCGGCGGCCACTTCTACCTCAACGACCAGTGGCACGAGATCTGCAACGACATCTCCGACCACCTGCTCG
GCAAGACGGAGGCCCGCATGAGGCCCGCGGTGAAGATGGAGTTGCTGGTCAACGCTCTAGACGTTGCTGTAGAGGCTGGTGGACGAGC
F C L R A Y S G G H F Y L N D Q W H E I C N D I S D H L L V

FIG. 32-2

sugar.finalgene b-1 Sequence

10	20	30	40	50	60	70	80	90
1234567890123456789012345678901234567890123456789012345678901234567890								
<div> <div>AsclI</div> <div>▼</div> <div>TCACCCGGCGCGCCCGATGCCCGCGTCTGTGACGCCCGACCGCCTTATCGAAGGAGCGCGGAAGAGATGGCAGAA</div> <div>AGTGGCGCGCGCGGCTACGGCGCGCAGCACGTGCGGGGCTGGTCGGAATAGCTTCTCGCCGCTTCTCTACCGTCTTTGGGTGCCACTG</div> <div>T R G A P D A R V V Q P P T S L I E G A A K R W Q N P R</div> <div>V T</div> </div>								
<div> <div>EarI</div> <div>▼</div> <div></div> <div></div> <div></div> </div>								
<div> <div>BstEII</div> <div>▼</div> <div></div> <div></div> <div></div> </div>								
<div> <div>810</div> <div></div> <div></div> <div></div> <div></div> </div>								
<div> <div>CGACGACCTGACGGGGGCCCTCACGACGCCCCCGCTGGGCCGACCGTCCGGCGGTGGCCGACCGTGAACTCGGCACCCACCTCCTGGA</div> <div>GCTGTGGACTGCCCCCGGGAGTGCCTCGGGGGCGACCCCGCGTGGCAGCGCGCCACCGCTGGCACTTGAGCCCGTGGGTGGAGGACCT</div> <div>D D L T G A L T Q P P L G R T V R A V A D R E L G T H L L E</div> </div>								
<div> <div>ApalI</div> <div>▼</div> <div></div> <div></div> <div></div> </div>								
<div> <div>PflMI</div> <div>▼</div> <div></div> <div></div> <div></div> </div>								
<div> <div>900</div> <div></div> <div></div> <div></div> <div></div> </div>								
<div> <div>BamHI</div> <div>▼</div> <div></div> <div></div> <div></div> </div>								
<div> <div>GACCCGGGCATCCACTGGATCCACGCCCGCGAACGGCGACCCCGTACGCCACCGTCTGCGCGGCCAGGCGGACGCCGTATCCCGCGTA</div> <div>CTGGCGCCGTAGGTGACCTAGTGCGGGCGCTGCCGTGGGCATCGCGTGGCAGCACGCGCCCGTCCGCCTGCTGGGCATAGGGCGGCAT</div> <div>T R G I H W I H A A N G D P Y A T V L R G Q A D D P Y P A Y</div> </div>								
<div> <div>BstEII</div> <div>▼</div> <div></div> <div></div> <div></div> </div>								
<div> <div>PvuII</div> <div>▼</div> <div></div> <div></div> <div></div> </div>								
<div> <div>CGAGCGGGTGCGTCCCGCGCGCGCTCTCCTTACGCCCGACGGGCAGCTGGGTCAACCGCCGATCACGCCCTGGCGGCGAGCATCCTCTG</div> <div>GCTCGCCACGACGGGCGCGCGGAGAGAGTCTGGGCTGCCGTCGACCCAGTGGCGGCTAGTGGCGGACCGCCGCTCGTAGGAGAC</div> <div>E R V R A R G A L S F S P T G S W V T A D H A L A S I L C</div> </div>								
<div> <div>1080</div> <div></div> <div></div> <div></div> <div></div> </div>								

FIG. 32 - 3

sugar.finalgene b-1 Sequence

10	20	30	40	50	60	70	80	90
1234567890123456789012345678901234567890123456789012345678901234567890								
CTCGACGGACTTCGGGGTCTCCGGCGCCGACGGGTCCGGTCCGACAGGTCCTCTCGTAGGGGAGGGGTGTCCGCTGGAGCGCGA								1170
GAGTGCCTGAAGCCCCAGAGGCCCGGGCTGCCGAGGGCCACGGCGTCTGTCAGGAGAGCATGCCCTCCCGACAGGCGACCTCGCGCT								
S T D F G V S G A D G V P V P Q Q V L S Y G E G C P L E R E								
<div> <div>AlwNI</div> <div>▼</div> </div> <div> <div>BamHI</div> <div>▼</div> </div>								
GCAGGTGTCGCCGGCGCGGTACGTGCCGGAGGGCGGCGAGCGTGCCTGTCGAGGGGATCCACGGGAGACGCTGGAGGGTCTCGC								1260
CGTCCACGACGGCCCGGCCACTGCACGGCCTCCCGCCCGTCCGACGGCACAGCTCCCTAGTGCCCTCTGCGACCTCCCGAGAGCG								
Q V L P A A G D V P E G G Q R A V V E G I H R E T L E G L A								
GCCGACCCGTCGGCGTCGTACGCCCTTCGAGCTGCTGGCGGTTTCGTCCGCCCGCGGTGACGGCCGCTGCCGCCGCCGTGCTGGGTGT								1350
CGGCCTGGGACGCCGACGATGCGGAAGCTCGACGACCCGCCAAGACGAGCGGGCCGCCACTGCCGGCGACGGCGGGCAGCACCCACA								
P D P S A S Y A F E L L G G F V R P A V T A A A A V L G V								
<div> <div>RsrII</div> <div>▼</div> </div> <div> <div>AlwNI</div> <div>▼</div> </div>								
TCCCGGGACCGGGCGCGGACTTCGCGGATCTGCTGGAGCGGCTCCGGCCGCTGCCGACAGCCTGTGGCCCCGAGTCCCTGCGGAC								1440
AGGGCGCCTGGCCCGCGCCTGAAGCGCCTAGACGACCTCGCCGAGGCCGCGACAGGCTGTCCGACGACCCGGGGCGTTCAGGGACGCCTG								
P A D R R A D F A D L L E R L R P L S D S L L A P Q S L R T								

FIG. 32 - 4

sugar.flnalgene b-1 Sequence

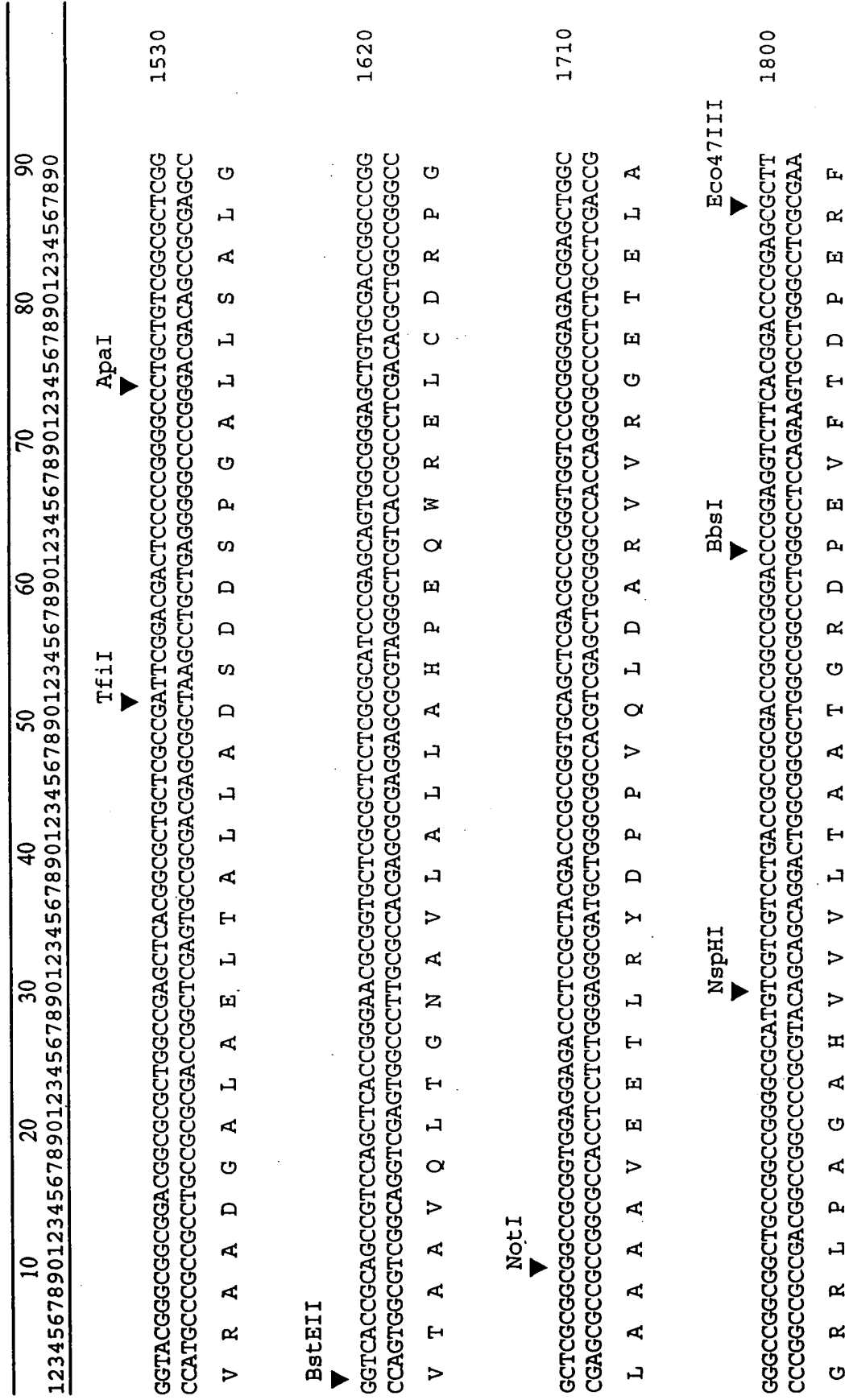


FIG. 32 - 5

sugar.finalgene b-1 Sequence

10	20	30	40	50	60	70	80	90
1234567890123456789012345678901234567890123456789012345678901234567890								
CGACCTCGGCGCCCGACCCCGCGGCACCTCGCGCTGCACCCCGCCGGTCCGTACGGCCCGGTGGCGTCCCTGGTCCGGCTTCAGGC								1890
GCTGGAGCGCGCGGGCTGCGGCGCGCGTGGAGCGCGACGTGGGGCGGCCAGGCATGCCGGGCCACCGCAGGGACCCAGGCCGAAGTCCG								
D L A R P D A A A H L A L H P A G P Y G P V A S L V R L Q A								
GGAGTTCGCGCTGCGGACCCCTGGCCGGGCGTTTCCCGGGCTGCGGCAGGGGGGACGTGCTCCGCCCCCGCGCGCCTGTGCGGCCG								1980
CCTCCAGCGGACGCTGGACCCGCCCGCCGCAAGGGCCCCGACGCCGTCCGCCCCCTGCACGAGCGGGGGCGCGCGGACAGCCCGG								
E V A L R T L A G R F P G L R Q A G D V L R P R R A P V G R								
CGGGCCGCTAGCGTCCCGTCAGCAGCTCCTGAGACACCGGGGCCCCGGTCCGCCCGGCCCCCTTCGGACGGACCGGACGGCTCGGAC								2070
GCCCCGCGACTCGCAGGGCCAGTCGTCGAGGACTCTGTGGCCCCCGGGGCCAGGCGGGCCCGGGGGGAAGCCTGCCTGGCCTGCCGAGCCTG								
G P L S V P V S S S								
CACGGGACGGCTCAGACCGTCCCGTGTGTCCCCGTCCCGCTCCCGCCCCCATCCCGCCCCCTCCACCGGCAAGGAAGGACACGACGC								2160
GTCCCCCTGCCGAGTCTGGCAGGGCACACAGGGGACAGGCCGAGGGCAGGGCGGGGAGGTGGCCGTTCTTCTGTGTCTGCTGG								

FIG. 32 - 6

sugar.finalgene b-1 Sequence

10	20	30	40	50	60	70	80	90
1234567890123456789012345678901234567890123456789012345678901234567890								
CATGCGGTCCTGCTGACCTCGTTCCGACATCACACGCACACTACTACGGCCCTGGTGCCCTTGGCCCTGGCGGCTGCTCGCCCGCCGGGCACGA								2250
GTACGGCAGGACGACTGGAGCAAGCGTGTAGTGTGCGTGATGATGCCGGACCAACGGGGACCGGACCCCGGACGAGCGGCGGCCCGCTGCT								
M R V L L T S F A H H T H Y Y G L V P L A W A L L A A G H E								
DraIII								
GGTGCGGTCGCCAGCCAGCCCGCTCACGGACACCATCACCGGTCGGGGCTCGCCGGTCCCGGTCGGGTCCGGCACCGACCATCTCATCCA								2340
CCACGCCCAGCGGTCCGTCCGGCGCGAGTGCCCTGTGGTAGTGGCCAGGCCCGAGCGGGCCACGGCCAGCCCGTGGCTGGTGGAGTAGGT								
V R V A S Q P A L T D T I T G S G L A A V P V G T D H L I H								
PvuI								
CGAGTACCGGGTGCGGATGGCGGGCGAGCCCGCCCCGAACCATCCGGCGATCGCCTTCGACGAGGCCCGTCCCGAGCCCGCTGGACTGGGA								2430
GCTCATGGCCCCACGCCTACCGCCCCGCTCGGCGCGGGCTTGGTAGGCCCGCTAGCGGAAGCTGTCTCCGGCAGGGCTCGGCGACCTGACCCCT								
E Y R V R M A G E P R P N H P A I A F D E A R P E P L D W D								
2520								
CCACGCCCTCGGCATCGAGGCGATCCTCGCCCCGTACTTCCATCTGCTCGCCAAACAACGACTCGATGGTCCGACGACCTCGTCGACTTCGC								
GGTGCGGAGCCGTAGCTCCGCTAGGAGCGGGGCATGAAGGTAGACGAGCGGTTGTGTGCTGAGCTACCAGCTGCTGGAGCAGCTGAAGCG								
H A L G I E A I L A P Y F H L L A N N D S M V D D L V D F A								

FIG. 32 - 7

sugar.finalgene b-1 Sequence

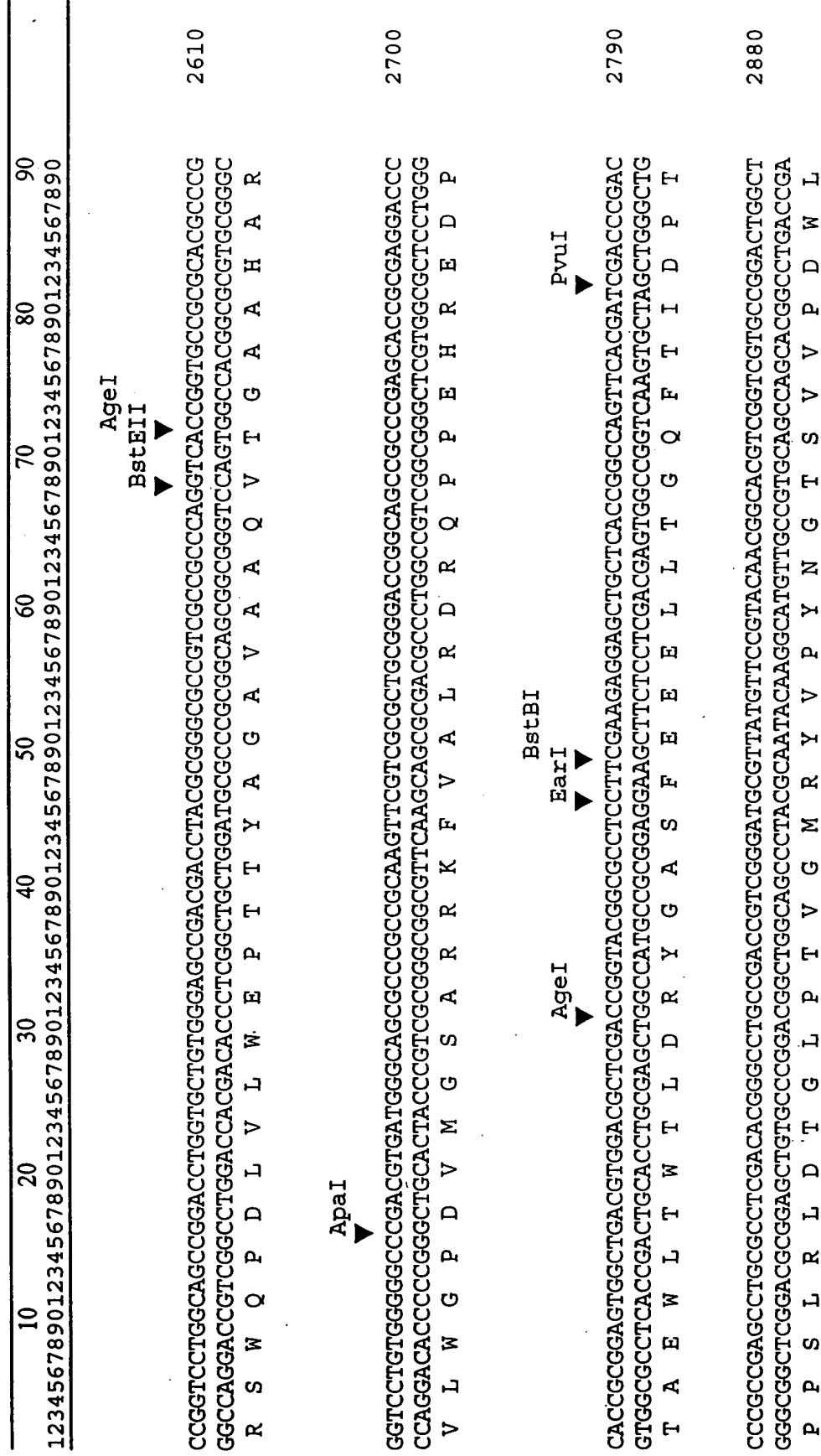


FIG. 32 - 8

sugar.finalgene b-1 Sequence

10	20	30	40	50	60	70	80	90
1234567890123456789012345678901234567890123456789012345678901234567890								
GAGTAGCCGCCCGCGGCCCGGGTCTGCCTGACCCCTCGGCGTCTCCGCGGTGAGGTCTCGGCGCGACGGCGTCTCGCAGGGCGGA								2970
CTCACTCGGCGGGCGGCCCGGGCCAGACGGACTGGGAGCCGCGAGAGCGGCACCTCCAGGAGCCCGCGTCCCGCAGAGCGTCCCGCT								
S E P P A R P R V C L T L G V S A R E V L G G D G V S Q G D								
CATCTTGAGGGCGTCCCGGACCTCGACATCGAGCTCGTCGCCACGCTCGACGCGAGTCAGCGCGCGAGATCCGCAACTACCCGAAGCA								3060
GTAGGACCTCCGCGAGCGGTGGAGCTGTAGTCTGAGCAGCAGCGGTGCGAGCTGCGCTCAGTCGCGCGGCTCTAGGCGTTGATGGGCTTCGT								
I L E A L A D L D I E L V A T L D A S Q R A E I R N Y P K H								
CACCCGGTTCAGGACTTCGTGCCGATGCACGCGCTCCTGCCGAGCTGCTCGGCGATCATCCACGCGGGCGGCACCTACGCGAC								3150
GTGGGCCAAGTGCCCTGAAGCACGGCTACGTGCGGAGGACGGCTCGACGAGCCGCTAGTAGTGTCGCGCCCGCCCGCGTGATGCGCTG								
T R F T D F V P M H A L L P S C S A I I H H G G A G T Y A T								
BclI								
CGCCGTGATCAACGCGGTGCCGAGGTATGCTGCCGAGCTGTGGACGCGCCGGTCAAGCGCGGGCCGTCGCCGAGCAGGGGGCGGG								3240
GCGGCACTAGTTGCGCCACGGCGTCCAGTACGAGCGGCTCGACACCCCTGCGCGGCCAGTTCCGGCGCCCGGCGAGCGGTCGTCCCCCGCCC								
A V I N A V P Q V M L A E L W D A P V K A R A V A E Q G A G								

FIG. 32 - 9

[illegible]

FIG. 32-10

sugar.final gene b-1 Sequence

10	20	30	40	50	60	70	80	90
1234567890123456789012345678901234567890123456789012345678901234567890								
<div> <div>StuI</div> <div>▼</div> </div>								
GAGGCCTCCGACATCGCCGACCTGGTGGCTCCGTCACCCCGAGGCCTCCTCGCTCCTGGACGTGGCCTGCGGTACGGGCACGCATCTG								3690
CTCCGGAGGCTGTAGCGGCTGGACACCGGAGGGCATGGGGCTCCGGAGGAGCGAGGACCTGACCGGACGCCATGCCCGTGGTAGAC								
E A S D I A D L V R S R T P E A S S L L D V A C G T G T H L								
<div> <div>StyI</div> <div>▼</div> </div>								
GAGCACTTCAACGAGGAGTTCGGCGACACCCCGGCTGGAGCTGTCCGAGGACATGCTACCCACGCCCGCAAGCGGCTGCCCGACGCC								3780
CTCGTGAAGTGGTTCCTCAAGCCGCTGTGGCGCGGACCTCGACAGGCTCCTGTACGAGTGGGTGGCGGCTTCGCCGACCGGCTGGCG								
E H F T K E F G D T A G L E L S E D M L T H A R K R L P D A								
<div> <div>NspHI</div> <div>▼</div> </div>								
ACGCTCCACGAGGACATCGCGGACTTCGGGCTCGGCCGGAAGTTCTCCGCCGTGGTCAGCATGTTTACGCTCCGTCCGTACCTGAAG								3870
TGCGAGGTGGTCCCGCTGTACGCCCTGAAGGCCGAGCCGGCCTTCAAGAGCGGCACCAAGTCGTACAAGTCGAGGCAGCCGATGGACTTC								
T L H Q G D M R D F R L G R K F S A V V S M F S S V G Y L K								
<div> <div>NspHI</div> <div>▼</div> </div>								
ACGACCGAGGAACTCGGCGGGCCGTGCCTCGTTCGGGAGCACCTGGAGCCCGGTGGCGTCTGTCGTCTGAGCCGTGGTGTCCCG								3960
TGCTGGCTCCTTGAGCCGCGCGGAGCAAGCGCCTCGTGGACCTCGGGCCACCGCAGCAGCAGCTCGGCACCCACCAAGGCG								
T T E E L G A A V A S F A E H L E P G V V V E P W F P								
<div> <div>BbsI</div> <div>▼</div> </div>								

FIG. 32 - 11

sugar.finalgene b-1 Sequence

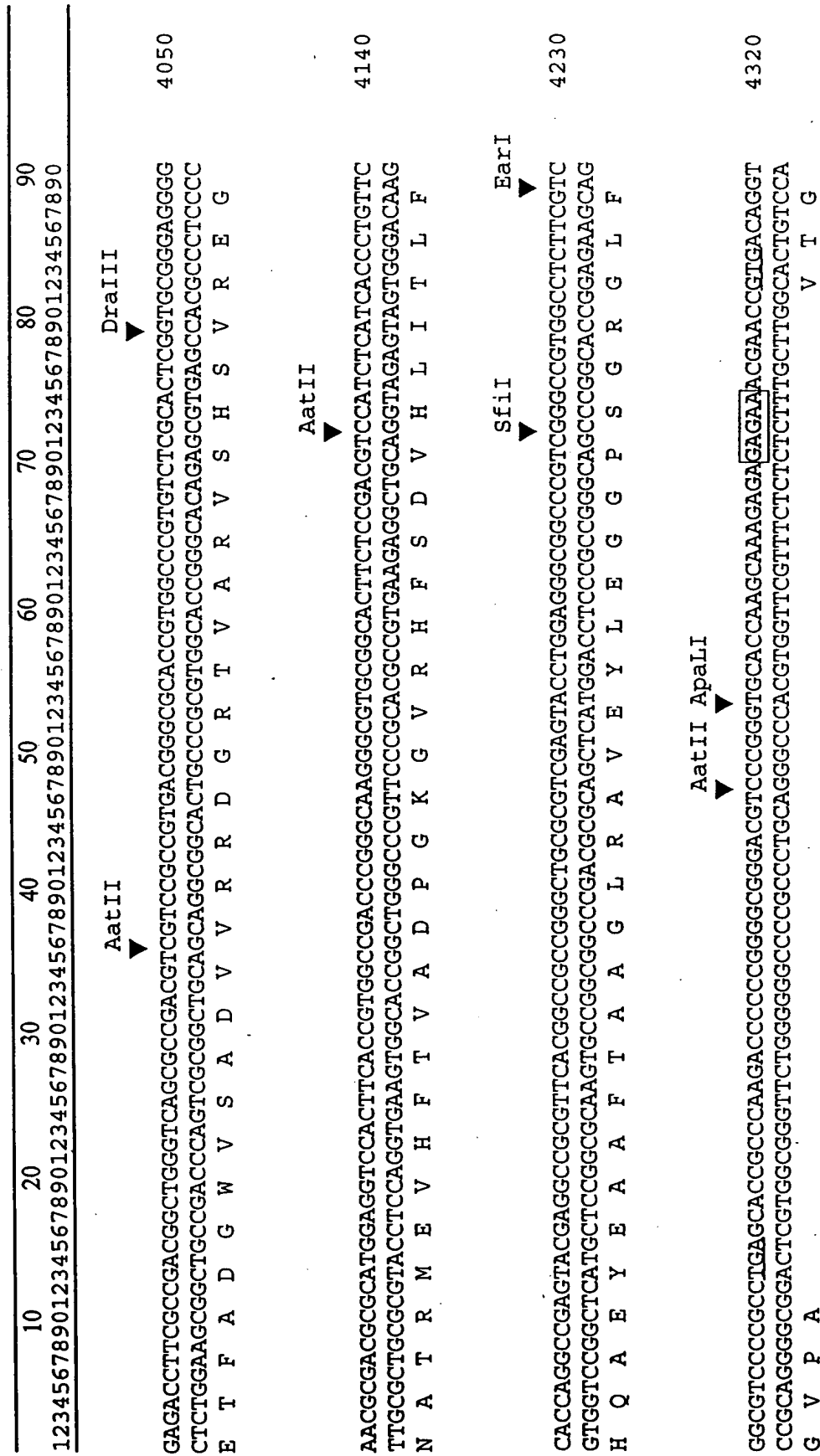


FIG. 32 - 12

sugar.finalgene b-1 Sequence

10	20	30	40	50	60	70	80	90
1234567890123456789012345678901234567890123456789012345678901234567890								
AAGACCCGAATACCGCGTGTCCGCGCGCGCCGACACAGCCCGGCGCTTACCCCTGGCCGTCGTCGGCACCCCTGCTGGCGGGCACCCACC								4410
TTCTGGGCTTATGGCGCACAGGCGGCGCGCGGTGGTGGGTCCCGGAAGTGGACCCGCGCAGCAGCCGTGGGACGACCGCCCGTGGTGG								
K T R I P R V R R G R T T P R A F T L A V V G T L L A G T T								
NotI								
▼								
GTGGCGCGCGCGTCCCGGCGCGCGCACAGGCCAATGTTCAGTACAGAGCCGGCGGGAGCTCGTCGCCAGATGACGCTCGAC								4500
CACCGCGCGCGAGGGCGCGCGGTGTGCCGTTACAAGTCAATGTGCTCGGCCCGCTCGAGCAGCGGGTCTACTGCGAGCTG								
V A A A P G A A D T A N V Q Y T S R A A E L V A Q M T L D								
NotI								
▼								
GAGAAGATCAGCTTCGTCCACTGGGCGCTGGACCCCGACCGGACAGTACGTCGGCTACCTTCCCGGCGTCCCGCTCTGGGCATCCCGGAG								4590
CTCTTCTAGTCGAAGCAGGTGACCCCGACCTGGGGCTGGCCGCTTTCAGCCCGATGGAAGGGCCGACGCGCGCAGACCCGTAGGGCCTC								
E K I S F V H W A L D P D R Q N V G Y L P G V P R L G I P E								
MscI								
Bali								
▼								
CTGCGTGCCCGACGGCCCGAACGGCATCCGCTGTGGGGCAGACCGCCACCGGCTGCCCGCGCCGTCGCCCTGGCCAGCACCTTC								4680
GACGCACGGCGGCTGCCGGCTTGGCGGACCCCGTCTGGCGGTGGCGGACGCGCGCGCCAGCGGGACCGGTCTGTGGAAG								
L R A A D G P N G I R L V G Q T A T A L P A P V A L A S T F								

FIG. 32 - 13

sugar.finalgene b-1 Sequence

10	20	30	40	50	60	70	80	90
1234567890123456789012345678901234567890123456789012345678901234567890								
<div> <div> <div>▼</div> <div>StyI</div> <div>NcoI</div> </div> <div> <div>▼</div> <div>PflMI</div> <div>Apal</div> </div> </div>								
GACGACACCATGGCCGACAGCTACGGCAAGGTCA	TGGGCCGCGACGGTCGGCGGCTCAAC	CAGGACATGGTCTCTGGGCCCCGATGATGAAC						4770
CTGCTGTGGTACCGGCTGTCGATGCCGCTTCCAGT	ACCCGGCGCTGCCAGCGCGAGTTGGTCTCTGTAC	CCAGGACCCCGGCTACTACTTG						
D D T M A D S Y G K V M G R D G R A L N Q	D M V L G P M N							
<div> <div> <div>▼</div> <div>StyI</div> <div>NcoI</div> </div> <div> <div>▼</div> <div>PflMI</div> <div>Apal</div> </div> </div>								
AACATCCGGGTGCCGCACGGCGCCGGAAC	TACGAGACCTTACGCGAGGACCCCTGGTCTCTCG	CGCACCCGGTCCCGCAGATCAAG						4860
TTGTAGGCCCCACGGCGTGCCCGCGCCTTGATGCT	CTGGAAATCGCTCTGGGGACCAAGAGGCGCGTG	CGCCAGCGGGTCTAGTTC						
N I R V P H G G R N Y E T F S E D P L V S	S S R T A V A Q I K							
<div> <div> <div>▼</div> <div>StyI</div> <div>NcoI</div> </div> <div> <div>▼</div> <div>PflMI</div> <div>Apal</div> </div> </div>								
GGCATCCAGGTGGGGTCTGATGACCA	CGGCCAAGCACTTCGGCGCCAAAC	ACCAGGAGAAACACCGCTTCTCCGTGAAC						4950
CCGTAGGTCCCAACGCCAGACTACTGGTGCCGGTT	CGTGAAGCCCGGTTGTTGGTCTCTTGTGG	CGAAGAGGCACCTTGGCGTTACAG						
G I Q G A G L M T T A K H F A A N N Q E	N N R F S V N A N V							
<div> <div> <div>▼</div> <div>StyI</div> <div>NcoI</div> </div> <div> <div>▼</div> <div>PflMI</div> <div>Apal</div> </div> </div>								
GACGAGACAGCTCCGCGAGATCGAGTTC	CCGGCGTTCGAGGCGTCTCCAAGCCGGCGG	CGCTTCATGTGTGCTACACGGC						5040
CTGCTCGTCTGCGAGGCGCTCTAGCTCAAGG	CGCGCAAGCTCCGAGGAGTTCCGGCCGCG	CGGAGGAGTACACACGGATGTTGCCG						
D E Q T L R E I E F P A F E A S S K A G	A A S F M C A Y N G							

FIG. 32 - 14

sugar.finalgene b-1 Sequence

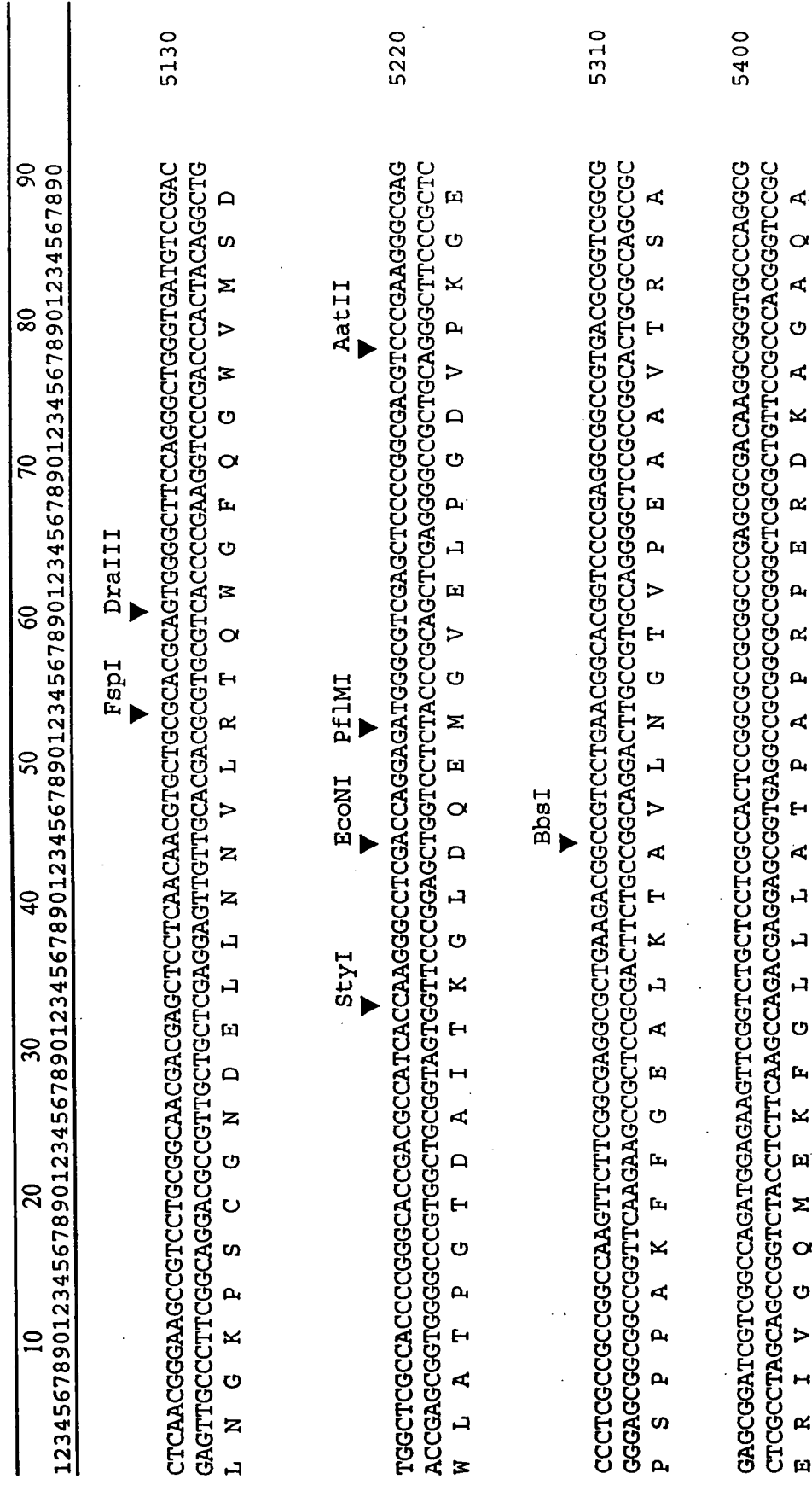


FIG. 32 - 15

[illegible]

AlwNI
EspI

BstEII
StyI ▼

XhoI
PaeR7I
▼

AGCCCCGGTTCAACCAAGGGCCACCAGCTCGAGCCGGGCAAGCGGGGGCGCTGTACGACGGCACGCTGACCGTGCCCCGCCGACGGCGAG
TCGGGCCGCAAGTTGGTCCCGGTGGTTCGAGCTCGGCCCGGTTCGGCCCCCGGCACATGCTGCCGTGCGACTGGCACGGGGCGGCTGCCGCTC
S P A F N O G H Q L E P G K A G A L Y D G T L T V P A D G E

FIG. 32-16

[illegible]

AGCAGCCCGCTCCTCAAGCTGACCAAGGGCAGCACCAAGCTCAGATCTCGGGCTTCGGCGATGAGTGCCACCCCGCTCTCCCTGGAGCTG
 TCGTCGGCGGAGGAGTTCGACTGGTTCCCGTGGTTCGAGTGTAGAGCCCGAAGCGCTACTCAGGTGGGCGGAGAGGACCTCGAC
 SSSPILIKLTKGLTKGTHKLTLTISGFAMSAATPLSLEL

NruI
 PvuI
 BbsI
 6030

GACGACGGCACCGAGGGCGTCCGACCGTCCGAACTGTGCTGCCGGGTACGCAGGACAAGCTGATCTCGGCTGTGCGGACGCCAACCCG
CTGCTGCCGTGGCTCCCGCAGCTGGCAGGCTTGGACAGCAGCGGCCCATGCGTCTGTTCGACTAGAGCCGACAGCGCTGCGGTTGGGC
P D G T E G V D R P N L S L P G T Q D K L I S A V A D A N P

FIG. 32-17

sugar.finalgene b-1 Sequence

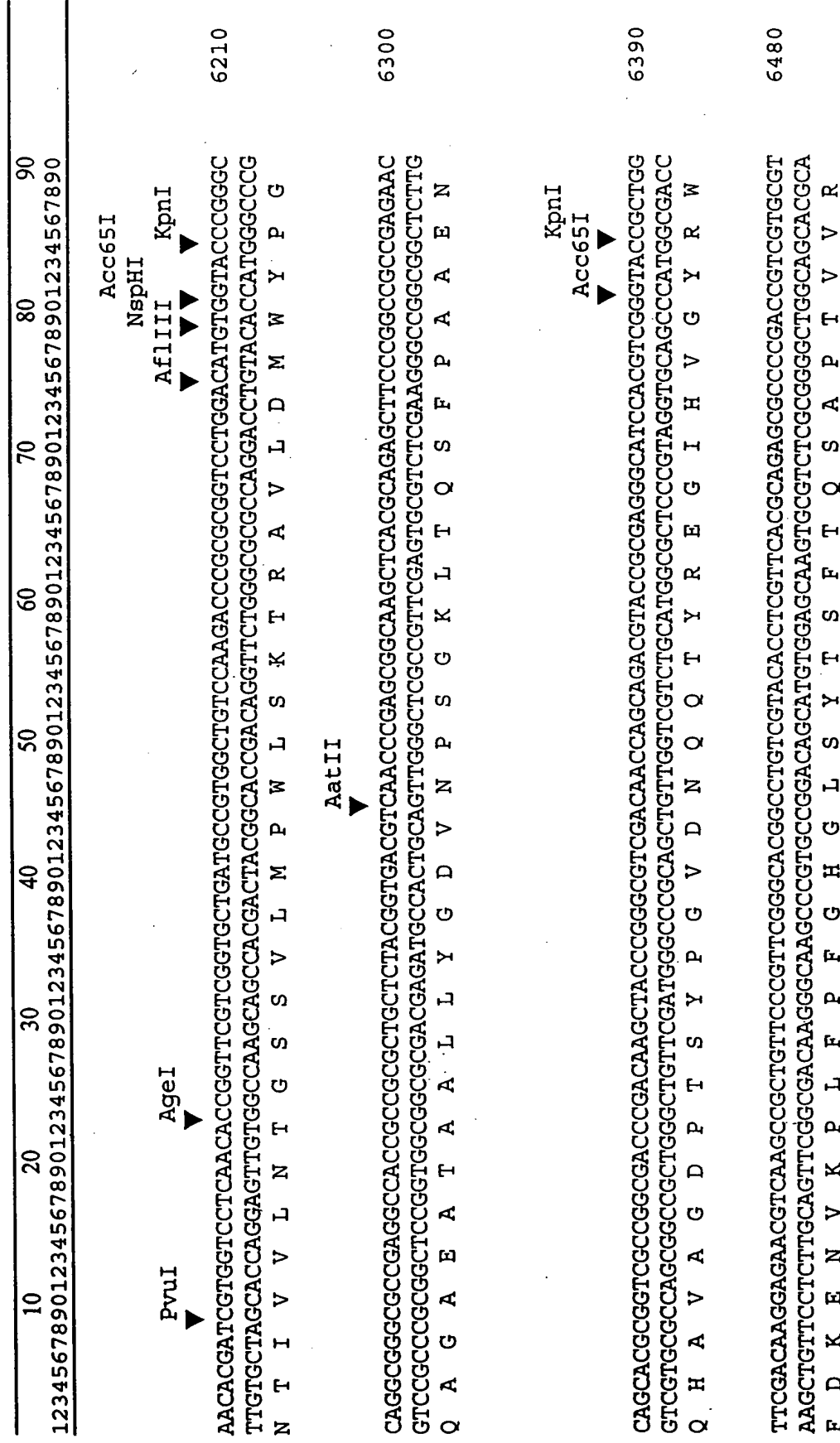


FIG. 32 - 18

sugar.finalgene b-1 Sequence

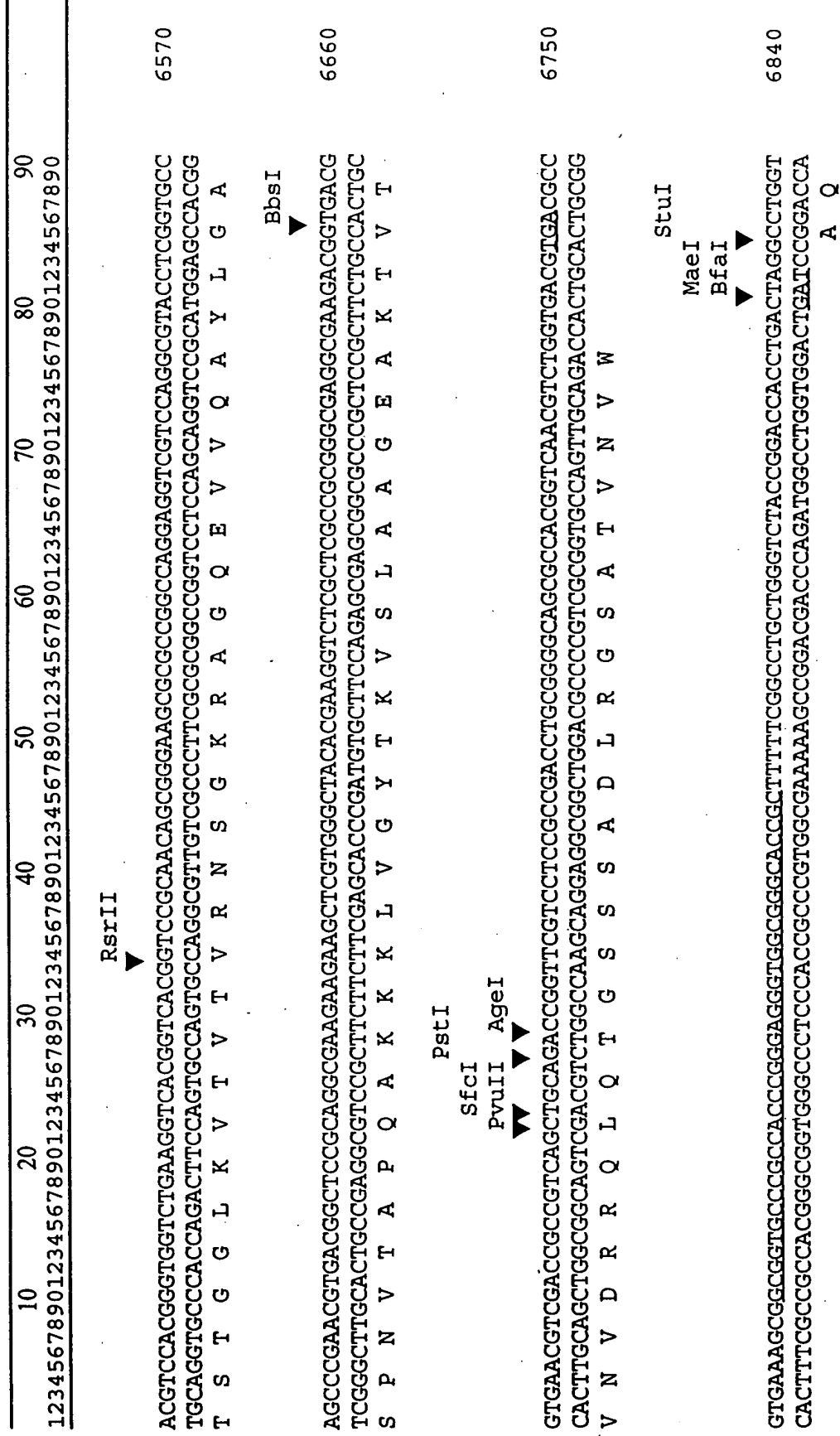


FIG. 32 - 19

[illegible]

FIG. 32-20

sugar.finalgene b-1 Sequence

10	20	30	40	50	60	70	80	90
1234567890123456789012345678901234567890123456789012345678901234567890								
PvuII ▼	BamHI ▼	XcmI ▼						
TCCAGCTGTCCAGGTGGCGAGCCGGATCCGCAGCACGGGGCCCTGCATCTCGTCCAGGGGAGTTGGTGCCTTCGTCTCGTGGCTGT								7290
AGGTCGACAGGTCCACCCGCTCGGCCTAGGCGTGTGCGCGGACGTAGAGCAGGTCCGCCCTCAACACAGGGAAGCAGAGCACCGACA								
N W S D L H A L R I R L V A A Q M E D L R S N T G K T E H S								
BspMII BspEI ▼								
ACTTCTCGCGGAGCCGTAGTTGCGGAGCATCCGGAGCCGTTTCGGGAGCTCGGGTTCGCCGTGACGACGGCGCCCGCTCGCCGAAGC								7380
TGAAGACGGCGCTCGGCATCAACGCCTCGTAGGCCTCGGCAAGCCGCTCGAGCCCCAGCGGCCACTGCTGCCGCGGCGCAGCGGCTCG								
Y K Q R S G Y N R L M R L R E A L E P D G T V V A G G D G F								
RsrII ▼								
AGCCGAGGTTCTTGCCCGGGTAGAAGCTGAACGGGGCCACCGACGACCCGGCGCCGATCCGCCGGCCCCCGGTAGCGGGCGCCGTGGGCCT								7470
TCGGCTCCAAGAACGGGGCCCATCTTCGACTTGGCCCGGTGGCTGCTGGGCCCGGCTAGCGGGCCCATGCCCGCGGACCCCGGA								
C G L N K G P Y F S F A A V S S G A G I R R G R Y R A G H A								
RsrII ▼								
GGCGGGCGTCTCGACGATGTGCAGGCCGTGCCGGTCCGGAGCTCGCGAGGGCGTCCATGTGGCGGGGTGCCGTAGAGGTGGACGG								7560
CGCGCCGAGGAGTGCTACAGTCCGGCACGGCCAGGCGCTCGAGCGCCTCCCGAGGTACAGCCGCCACCGGCATCTCCACCTGCC								
Q A A D E V I H L G H R D A L E R L A D M D A P H G Y L H V								

FIG. 32 - 21

sugar.finalgene b-1 Sequence

10 20 30 40 50 60 70 80 90
1234567890123456789012345678901234567890123456789012345678901234567890

GGAGGAGCGCCCGGGGTGCGGGGGGTGATCGCCCTTCTCGACGAGACGGGGTCCAGGGTGGGTGGTCTCTGTGGGCTCGACGGGCACGG 7650
CCTCCTCGGGGGCCACGCCCCCCTAGCGGAAGAGCTGCTCGTGGCCAGGTCCACCCACAGGAGCACGCCGAGCTGCCCGTGCC
P L L A R T R P T I A K E V L L P D L T P H D E H P E V P V

BsaAI
PvuII AflIII

GGGTGCGCGCGGTGGCGACACCGGAGCCAGCTGGCGATGTACGTGTCGAGGGGACGATCACTCGTCCCCGGGTCCGATGCCGAGGC 7740
CCCAGCGGGCCACCGCCTGTGGCGCTCGGTGACCGCTACATGCACACGCTCCCTGCTAGTGAGCAGGGGCCAGGCTACGGCTCCG
P T A G T A S V A L W S A I Y T H S P V I V E D G P G I G L

TfiI

CGCGAGGGCGAGCTGAGGGCGTCCATCCCGTGTTCACGCCGACGGCGTGGTCCGTCTCGCAGTACGCGGCGAACTCCGCCCTCGAATC 7830
GCGCCTCCCGCTCGACCTCCCGCAGGTAGGGCGACAAGTGGCGTCCCGCACAGGACAGCGTCAATGCGCCGCTTGAGGGGAGCTTAG
G R L A L Q L A D M G S N V G V A H D T E C Y A A F E A E F

PvuI

CTTCGAGTTCGGGTCCGAGGAGGTAGCGCCCCCGAGTCGAGGACGCGGGCGGATCGCGCGTCTCGGCGGAGCTCCTCGTAGGCGG 7920
GAAGCTCAAGCCCAGGCTCCTCCATCGCGGGGCTCAGCTCCTGCGCCCGCTAGCGCCGCGCAGCGCGGCGCTCGAGGAGCATCCGCC
G E L E P G L L Y R G S D L V R A I A A D T E A R L E E Y A

FIG. 32 - 22

sugar.finalgene b-1 Sequence

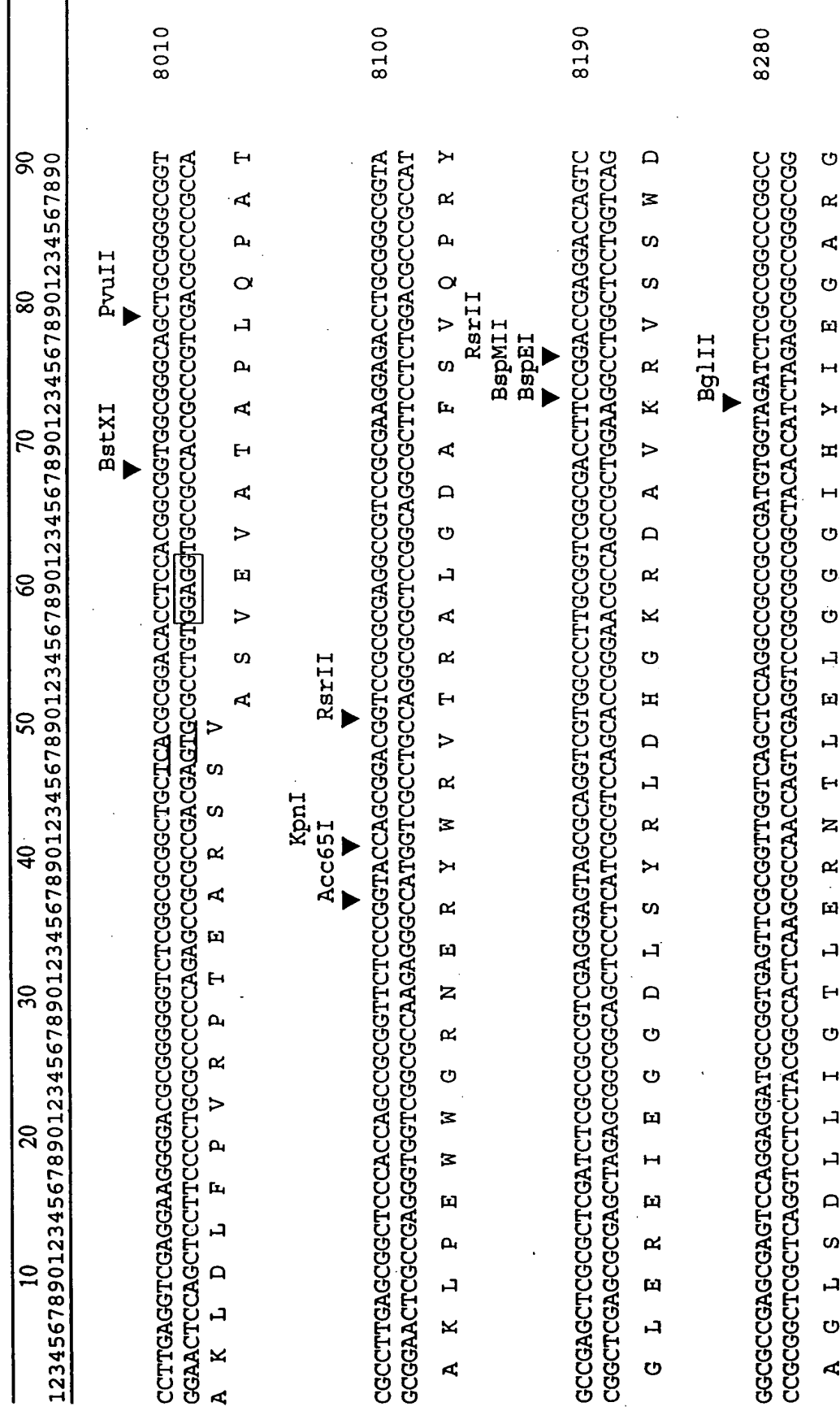


FIG. 32 - 23

sugar.finalgene b-1 Sequence

10	20	30	40	50	60	70	80	90
1234567890123456789012345678901234567890123456789012345678901234567890								
ApaLI ▼								
GCCCGGAGGACGAGCGGATGCCCGGACAGTGGTGGTGTGACACCCACTCGCGGACGTTGGCGCGCTGCCGTACAGCGGGAGCGGT 8370 CGGGCGCTCCTGCTCGCGCTACGGGGCCGCTCACGAGCAGCCACACAGTGGGTAGCGCCTGCAAGCGCGGACGCGGCAATGTCGCCCTCGCA								
G A L V L A I G R C H D D T H V W E R V N A G D G Y L P L T								
EarI ▼								
CCCCCGTCGAGGAGGTTCGTCACGAAGAGGGGATGAGCTTCTCGGGTGCTGGTACGGCCCGTAGTTGTTGTCAGACGCGGTGATCCG 8460 GGCGGCAGCTCCTCCAAGCAGTGTCTTCCCCCTACTCGAAGAGCCCCACGACCATGCCGGGCATCAACAACGTGCTGCCCCACTAGGC								
G G D L L N T V F L P I L K E P H Q Y P G Y N N C C R T I R								
StyI ▼								
TAGTCGAGGCCGTACGTCCGGTGGTAGGGCGGGCAACGAGGTCGGAGCCGGCCCTTGGACGCCCGGTAGGGCGAGTTGGGCTCCAGCGG 8550 ATGAGCTCCGGCATGAGGCCACCATCCGGGCCCGTTGCTCCAGCCTCGGCCGGAACCTGGCGGCATCCCGCTCAACCCGAGGTGCCC								
V D L G Y T R H Y A R A V L D S G A K S A A Y P S N P E L P								
PvuI ▼								
GCTGCTCGGTCCAGGAGCCGGAGTCGATCGACCCCGTACACCTCGTGGTGGAGACGTGCAACGACCCCGCCGACGCCGGCGTTCGACGGC 8640 CGACGAGAGCCAGGTCCTCGGCCCTCAGCTAGCTGGGCATGTGGAGCAGCCACCTCTGCACGTGCTGGCCCGGCTGCGGCCGACGCTGCCG								
S S E T W S G S D I S G Y V E D T S V H V V R G V G A D V A								

FIG. 32 - 24

sugar.final gene b-1 Sequence

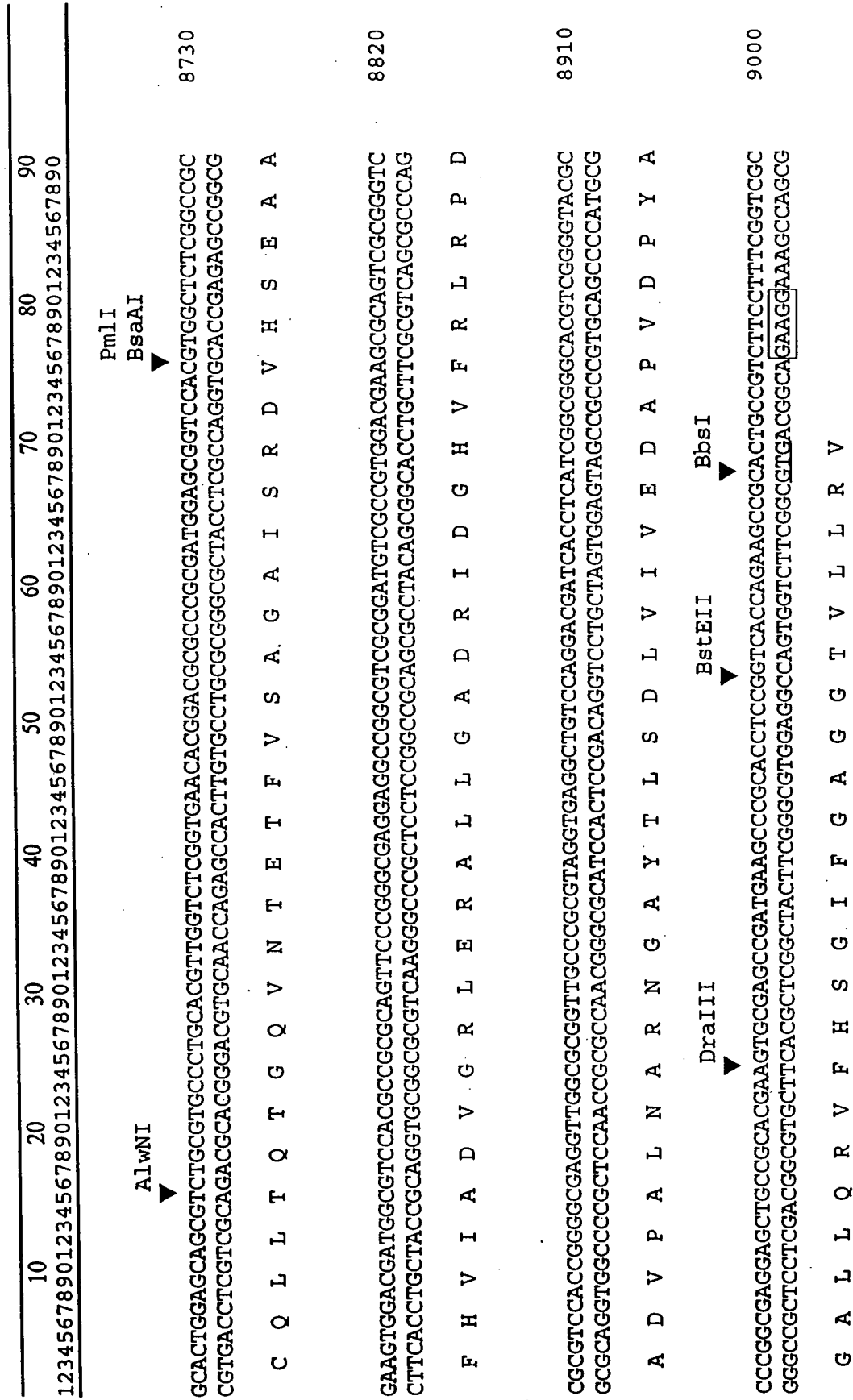


FIG. 32 - 25

[illegible]

SfcI
 MluI
 AflIII
 9090

BsaBI StuI
 ▼ ▼
 GCGTCGGCCGCGAGTGCCCTCACGGGGCTCCCTCGGGCCGGCCATCTCCATCAGATAGTGCCTACTCGGTCCGGAGAGGCCCTTCT
 CGCAGCCGGCGCTCCACGGGAGTGCCTCCGAGGGAGCGCCGGCCCTAGAGGTAGTCTATCGACGGCATGAGCCACGCCCTCTCCGGAAGA
 P A G E R G A I E M L Y S G Y E T R S L G E

StuI
▼
CCCAGGCCGTACAGGCCTCGGCGTCGATGAAGCCCATGCGGAAGGCGATCTCTCAAGCCCGCGATCCAGACGCCCTGCCGCTCCTCC
GGGTCCGGCACTGTCCGGAGCCGCAGCTACTTCGGGTACGCTTCGCTAGAGGAGTTCCGGGCGCTAGGTCTGCGGACGCCGAGGAGG
G L G H C A E A D I F G M R F A I E E L G A I W V G Q R E E

Pf1MI
AlwNI
▼
AGGACCTGGACGTACTGGGCGGCCCCGACGAGCGAGTGGTGCGTGGTGCCGGTGTCAGCCAGCGGAAGCCGGGCCACAGTTGACGAGTTCG
TCCTGGACCTGCATGACCCCGCGGGCGTCCCTCGCTCAGCACCCACGCGCCACAGTTCGGTCCGCTTCGGCGCCGGGTCCCAACTGCTCAAGC
L V Q V Y Q A A R L L S D H T G T D L W A F G R G L N V L E

FIG. 32 - 26

sugar.finalgene b-1 Sequence

10	20	30	40	50	60	70	80	90
1234567890123456789012345678901234567890123456789012345678901234567890								

AatII



GCCCCCCCCCTCCAGGTAGACGCGGTTGACGTCGGTGATCTCCAGCTCGCCGCGGCGAGGGCCGGATGTTCTTGGCGATGTGACG 9450
 CGGCGCGGGCGAGGTCCATCTGCGCCAACTGCAGCCACTAGAGTTCGAGCGGCGCGCCGCTCCCGCCTACAAGAACCGCTACAGCTGC
 A R G R E L Y V R N V D T I E L E G R P S P R I N K A I D V

AatII



ACGTCGTTGTCGTAGAGGTAGAGCCGGTGACGGCGAGGTTGGAGCGCGGCTTGACGGGCTTCTCGACGAGGTGGTCAGCCGGCCCGTC 9540
 TGCAGCAACAGCATCTCCATCTCCGGCCACTGCCGCTCCAACCTCGCGCCGAAGTGCCTCAGCCAGTCCGCGCGGCGAG
 V D N D Y L Y L G T V A L N S R P K V P K E V L D T L R G T

EarI



GCGTCCACCTCGGCGACGCCGTACCGCTCGGGGTCTTGACCGGGTAGCCGAAGACACGACGCCGTCGAGGCGCGGATGCTGTCCCGC 9630
 CGCAGGTGGAGCCGCTGCGGCATGGCGAGCCCCAGGAAGTGGCCCATCGGCTTCTCGTGGTTCGGCAGCTCCGCGCGCTACGACAGGGCG
 A D V E A V G Y R E P D K V P Y G F L V C G D L R A I S D R

ApaI



AGGAGCGGTAGAGCGCGGGCCCGTGGAGATGTTGTGCCCAGGATCAGGGCGCAGGTGTGTCGCCGATGTGTCGGCTCCGACGAGA 9720
 TCCTCGCACATCTCCGGCCCCGGCACCTTCTACAACAGCGGTCTTAGTCCCGGCTCCACAGCAGCGGCTACACGAGCCGAGGCTGCTCT
 L L T Y L G P G H F I N D G L I L A C T D D G I H E A G V L

FIG. 32 - 27

sugar.finalgene b-1 Sequence

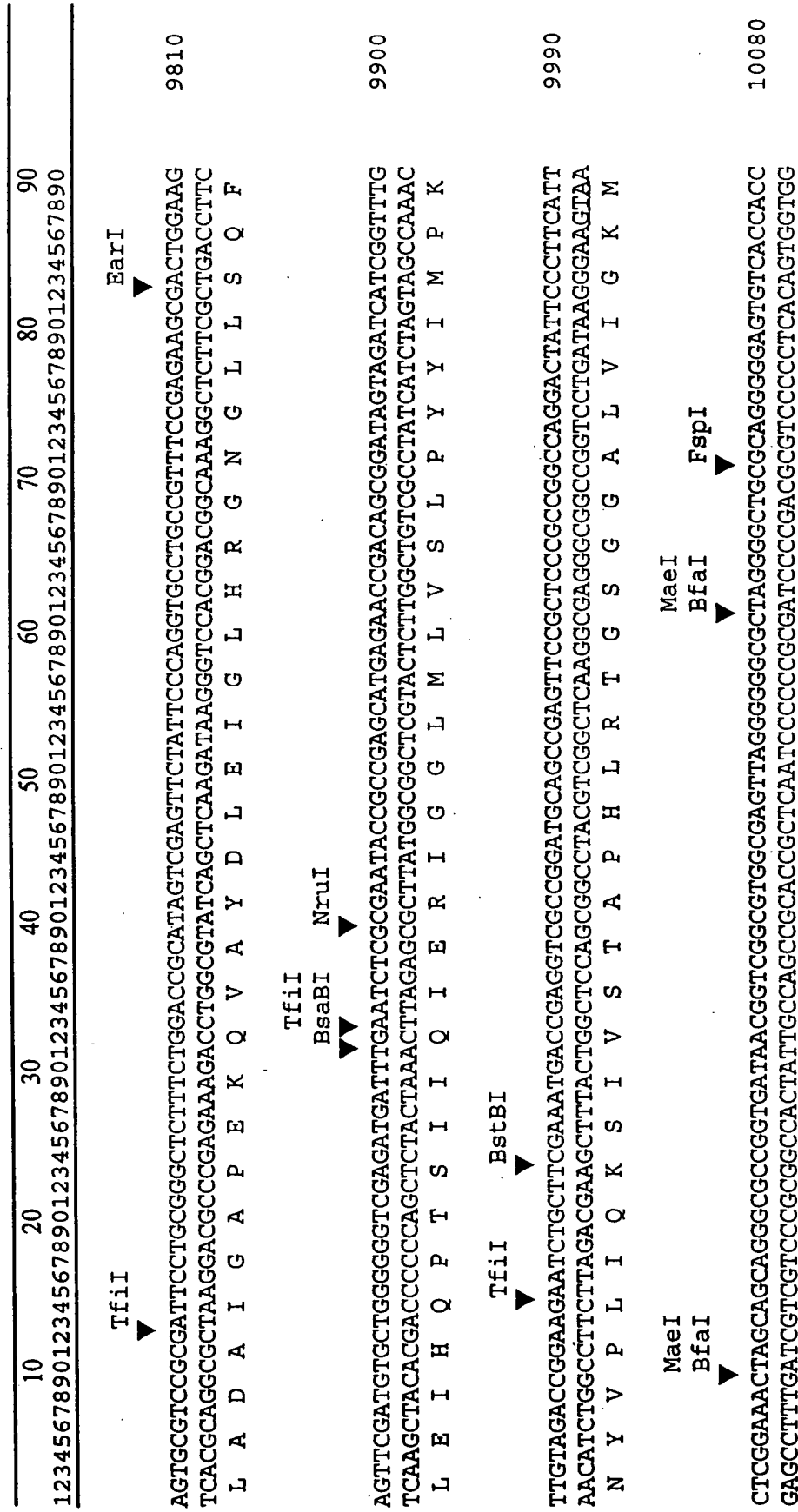


FIG. 32 - 28

[illegible]

FIG. 32 - 29

sugar.finalgene b-1 Sequence

10	20	30	40	50	60	70	80	90
1234567890123456789012345678901234567890123456789012345678901234567890								
AGGAGATCGACCTGCACCGGACCTGGGGGTGCGGGTGGGCCGCGATGGTGGCGGGCTTGATCCGCAGCAGTTCCGGCGTCCGCCCGCGGTG								10530
TCCTCTAGCTGGACGTGGCGCTGGACGCCACCGCCACCGCGGTACCAACGCCCGAACTAGGCGTGTCAAGCCGAGCCCGGGGCCAC								
L L D V Q V A V Q P H A T P R M T A P K I R L L E A D A G T								
FspI								
CGCAGGCTGTTCAAGGGCGTAGCCGTAGTCGATGTGGAGTCCGGGGGTGCGGTCCGGGACCCCGCTCCTCGAAGGCGTTGAGGGCCTCCTGG								10620
GCGTCCGACAAAGTCCGCGATCGGCATCAGCTACACCTCAGGCCCCCGACGCGAGCGCCTGGCGGAGGAGCTTCCGCAACTCCCCGGAGGACC								
R L S N L A Y G Y D I H L G P T R E R V R E E F A N L A E Q								
AGCTCGGCCCCGCTCCTCGGGCAGCTTGCCGTCGTCACGGCCGCTGTAGTCTCTCGGAATGTTGACGAAGTCGATCGTCTGCCCTGC								10710
TCGAGCCGGCGAGGAGGACGCCGTCGAACGGCAGCAGTCCCGCGACATCAGGAGCGCTTACAACCTGCTTCAGCTAGCAGGACGGGACG								
L E A R E E Q P L K G D D R G S Y D E R I N V F D I T R G Q								
CCGGCGTCGTTAGGTCGGCGATGAAGTCGACCAAGTCGAGCGGAGGACGGCCCGGAGCAGATGTAGCGAAGCCGAGGTTG								10800
GGCCGCAAGCACTCCAGCCGCTACTTCAGCTGGTCCAGCTCGTCCGCCCTCCGTCGGGCCCTCGTGTACATCCGCTTCGGCTCCAAC								
G A D N L D A I F D V L D L L R S A R G P L V I Y A F G L N								

FIG. 32 - 30

sugar.finalgene b-1 Sequence

10	20	30	40	50	60	70	80	90
1234567890123456789012345678901234567890123456789012345678901234567890								

AlwNI

PvuII

▼

ATCGGCGACTCGCGCTCGGCGCGCAGCTGCTGGAAGCGCGCAGGTTCTCGCGACGCGGCGGAAGGGCGCTTCTTCCGGTGGTCTGC
TAGCCGCTGAGCGGAGCCGCGCTCGACGACCTTCGCCGCGTCCAAGAGCGCTGCGCCGCTTCCGCCGGAAGAACGGCCACCGACG
I P S E R E A R L Q Q F R R L N E R V R R F A A K K G T T Q 10890

SfiI

▼

TCGTACTCCTCGTCTGAGGCCGTAGAGCGAGGTGCGGATGGGTGAGGCCCGCAGAGGCCGGCTGGCGCTCCAGGGTGGCTCGGTG
AGCATGAGGAGCAGCAACTCCGGCATCTCGCTCCACGCCTACCGCACGTCCGGGTCTCCGGCCCGACCGGAGGTCCCAACGGAGCCAC
E Y E E D N L G Y L S T R I A H L G W L G P Q R E L T R E T 10980

XmnI

▼

AGCGGAAGGAGTTCGTGTAGACGGTGGCGCCGAGGCCGTGGTGGGTGCGCGGCCAGGCTCCCGAGGCCGGGTGTGTGAGCGGC
TCGCGCTTCCTCAAGCACATCTGCCACCCGCGTCCGGCACCGACCCGCGCGGTCCGAGGGCTCCGGCCCCCAACCACTCGCCG
L A F S N T Y V T P R L G H D T A H A A L S G L G P N T L P 11070

TCCAGGCCCGCGGAGAAGTACATCGCCGAGGGGTGCCCGCGGGTATCTCGTCGATACCGACCGGAACATGGCGTTCGCGGCGTCGAGG
AGGTCGGCGGCGCTTTCATGTAGCGGCTCCCCAACGGGCGCCCATAGAGCAGCTACTGGCTGGCTTGTACCGCAACGGCCGCGAGCTCC
E L G G S F Y M A S P N G A P I E D I V S R F M A N G A D L 11160

FIG. 32 - 31

sugar.finalgene b-1 Sequence

10	20	30	40	50	60	70	80	90
1234567890123456789012345678901234567890123456789012345678901234567890								

NspHI



GCGACGGGTCGTAGCGGGCGCCGGTCACACGGACGCAGAAAGTGCGAGCGGAAACATGACAGGTCGGGCCGGGGTAGAGCCGACGCTGTAC
CGCCTGCCAGCATCGCCCGGGCCAGTGTGCCTGCGTCTTACCGTCGCCCTTGACGTCCAGCCCGGCCCAATCTCCGGCTGCGACATG
A S P D Y R A G T V R V C F H C R F M C T P G P Y L G V S Y 11250

BbsI



BbsI



GGGAAGACGGGCTTCTGCGAGCGCCGCGTCTGAAGACGCCCGCGCTGTCGAGCGGAGCAGGGTGTCTTCCAGTACGCCCCGGCGGGG
CCCTTCTGCCCCGAAGACCGCTCGCGGGCGCAGCTTCTGCGGCGCGACAAGCTCGCCCTCGTCCCAAGAAAGGTCTATGCGGGGCGCCCC
P F V P K R A L A A D F V G R Q E L P L L T N K W Y A G A P 11340

PflMI



CCGGTCTCGACCGCGGTGCGGAGCTCCGGGACCTGCCCGAACAGGGCGAGAGGCGCGGAAGCGTCCCGGTCCAGCCCGAGGTCTGTGG
GGCCAGAGCTGGCGCCACGCCTCGAGGCCCTGGACGGGCTTGTCGCGCTCCTCCGCGCCCTTCGCGAGGGCCAGCTGCGGGTCCAGCACC
G T E V A T R L E P V Q G F L A L L R R F A D R D V G L D H 11430

CGGGCTCTCCAGCGGGGTGAAGGGGCTGTGCCGTAGCGCACGGCGAGCCCGACGAGGTGGCGGGCGGGTCTCGGCCCTCGTCGGGC
GCCCCGAGGAGTGGCCCACTTCCCCGACAAACGGCATCGCGTGCCGCTCGGCCCTGCTCCACCGCCCCCAGCAAGGCCGAGCAGCCCCG
R A E E L P T F P S N G Y R V A L R V L H R A T T G A E D P 11520

FIG. 32 - 32

[illegible]

GCAGTGGCGGAAAGGCGGGCGGTTCATCGGGAGCGTCCAAATCGTGGGCGTGGATGTCTGGGGGGCCGCCAGCGGGGGCGCGTGT
CGCCACCGCCTTCCCGCCCCCGCCAGTAGCCCTCGCAGGTTAGCACCCGCCACCTACAGACCCCCCGCGCTCGCCCGCCCCCGGCACA
A T A S L A P A T M R S S R G I T P T S T O P A A L P A P A T

▶

CGCGGTGGCGCGGGTCA GTTCGGCGCGCGGGTCCGGCAGAGACGCAGCAGGTCCGGCGACCGCGCGGATCTCGTCTCGTCCGCGGATGGCGG
GGGCCACCGCGCGCCAGTCAAGCGCGCGGCCACGCGGTCTCTCGTCTCGTCCAGCGCTGGCGCGCCTACAGCAGCAGCGGCTACCGCC

DRHRATLERGRATACLLRLDLDAVRRIRDDDDGIA

TGCCGGTCCGACAGGACAGCAACGGCGCGCGGCGAGGGGTTCCGTTGTCCGGCAGCGGGGGCGTCCGGCTGCCCGCGGTACGGGCTCCAGCTCGT
ACGGGCAAGCCGTCCCTGTCTGTGCGCGCGCGCGCTCCGCAAGCCACACGCCGTCCGCCCGCACGCCGACGGGGCGCATGCCGAGGTCGAGCA

T G T P L S L V R A A L R E T H P L P A H P Q G R Y P E L E

FIG. 32-33

sugar.finalgene b-1 Sequence

10	20	30	40	50	60	70	80	90
1234567890123456789012345678901234567890123456789012345678901234567890								
<div> <div>▼</div> <div>▼</div> <div>▼</div> <div>▼</div> <div>▼</div> <div>▼</div> <div>▼</div> <div>▼</div> <div>▼</div> </div>								
GGCAGCCCGCGAGAAGTAGGCGCGGGTGTGCACGCCCTTCGGCCTTCAGGACCTCCATACGAGGTGCGGGTGGATGCCGGTGGTGGCCT								11970
CCGTCCGGCCCGCTCTTCATCCGGCGCCACACGTGCGGAAGCCGGAAGTCTTGAGGTACTGCTCCAGCGCCACCTACGGCCACCAACCGGA								
H C G P S F Y A R T H V G E A K L V E M V L D R H I G T T A								
<div> <div>▼</div> <div>▼</div> <div>▼</div> <div>▼</div> <div>▼</div> <div>▼</div> <div>▼</div> <div>▼</div> <div>▼</div> </div>								
CGTCGATCTCGACGATCACGTA CTGGTGGTGTGAGGCCGTGGCGGTGCTGGTGGCGACGAGGACGCCGGGAGGTCCGCGAGGTGCT								12060
GCAGTAGAGCTGTAGTCATGACCACCAACAACCTCCGGCACCGCCAGCACCGCTGCTCTGCGGCCCTCCAGGGCGTCCACGA								
E D I E V I V Y Q H N N L G H R D H D A V L V G P L D A L H								
<div> <div>▼</div> <div>▼</div> <div>▼</div> <div>▼</div> <div>▼</div> <div>▼</div> <div>▼</div> <div>▼</div> <div>▼</div> </div>								
CGCGGTAGGCGGCGTGGTTGCGCCGGTTCGGTTCGATGACCTCGGGAACCGCTCGAGGGAGGTGAGGCCCATGGCGGCGCGGCTCGC								12150
GCGCCATCCGCGCACCAACGCGGCCAAGGCCAGCTACTGGAGCCCTTTGCGCAGCTCCCTCCACTCCGGGTACCGCCGCCGCCGAGCG								
E R Y A A H N R R N R D I V E P F A D L S T L G M A A A E								
<div> <div>▼</div> <div>▼</div> <div>▼</div> <div>▼</div> <div>▼</div> <div>▼</div> <div>▼</div> <div>▼</div> <div>▼</div> </div>								
TCATCTTGGCGTTGGTCCCGCCGGCGGGGTGCCCGCGGCAGGTGGAAGCCGAAGTTGTGAGGGCGCGGATCCGGCGCGGAGGTGG								12240
AGTAGAACCGCAACAGGGCGGCGCCCGACGGCGGCCCGTCCAGCTTCGGCTTCAACACCTCCCGCGCTAGGCCCGCGCTCCAGCC								
S M K A N T G G A P S G G P L D F G F N H L A R I R A A L D								

FIG. 32 - 34

[illegible]

Isga

PvuII

▶

CGGCGGCGCAGGGGCGGCCACAGAGGTGACGCCGACGACGGCGGTGTGACCGCGGCGGCCACCTGTGTCGGGTTCGAGGT
GCGCCGCGTCCCGCGGGGTCTCCACTGTGGCTGTGCGGCTCCACGCCCCCACTGTGGCGCCCGGTGGACCAAGGCCACGCTCCA

IsgB

FIG. 32 - 35

sugar.finalgene b-1 Sequence

10	20	30	40	50	60	70	80	90
1234567890123456789012345678901234567890123456789012345678901234567890								
<div> <div>BclI</div> <div>▼</div> </div>								
<div> <div> <div>MscI</div> <div>Bali</div> <div>▼</div> </div> <div> <div>SphI</div> <div>NspHI</div> <div>▼</div> </div> </div>								
TGATCACTTCGCCGGTGAGCGCGCGCGGTGCGCGAGGAGCTGGAGCCCGCGCGTGGCGTTGCAGGTGGCCACGGCATGCCGGACCCCGG								12690
ACTAGTGAAGCGGCCACTCCGGCCCGCGCACGGCTCCTCGACCTCGGGCCGGCACCGCAACGTCACCGGTGCCGTACGGCCTGGGGCC								
<div> <div>AlwNI</div> <div>▼</div> </div>								
CGAGCCCGGACGCGCTCCTCGAACTCGCGGACGAGCGGGCCCGCGTGGACAGCCACTGGCTGTCGAGGGCCCGGTGAGCCGCTCGT								12780
GCTCGGGCCCGCTGCGCGGAGGACTTGAGCGCCTGCTCGCCCGCGGCAACCTGTCGGTGACCGACAGCTCCCGGGCCAGCTCGGGCGAGCA								
<div> <div>BsmI</div> <div>▼</div> </div>								
ACAGCTGGCGCGGTGATGCGGTTGGGCGCGCCACAGGAGCGGCTGGTCGAAAGCGGGGGGGCCCGAAGATGCCAGGTCGGATA								12870
TGTCGGACCGCGCCAGCTACGCCAACCCGGCGGGTGCTCTCTCGCCGACAGCTTTCGCCCGCCCGCGGGCTTCTTACGCTCCAGCCTAT								
<div> <div>TfII</div> <div>XmnI</div> <div>▼</div> </div>								
Y L R A R D I R N P R G V L L P Q D F A A P G G F A L D S								
<div> <div>TfII</div> <div>▼</div> </div>								
AGGCGCTTTTACGGATGTTCCCTCCGGGGCCACCGTCACGAAATGATTGCCCGATCCGGGAATCCCGAACGAGGTGCCCGGCTCCACCG								12960
TCCGCGAAAGTGCCCTACAAGGGAGGCCCCGGTGGCAGTGCCTTTACTAAGCGGCTAGGCCCTTAGGGCTTGCTCCAGCGGGCGGAGGTGGC								
L A S K V								

FIG. 32 -- 36

sugar.finalgene b-1 Sequence

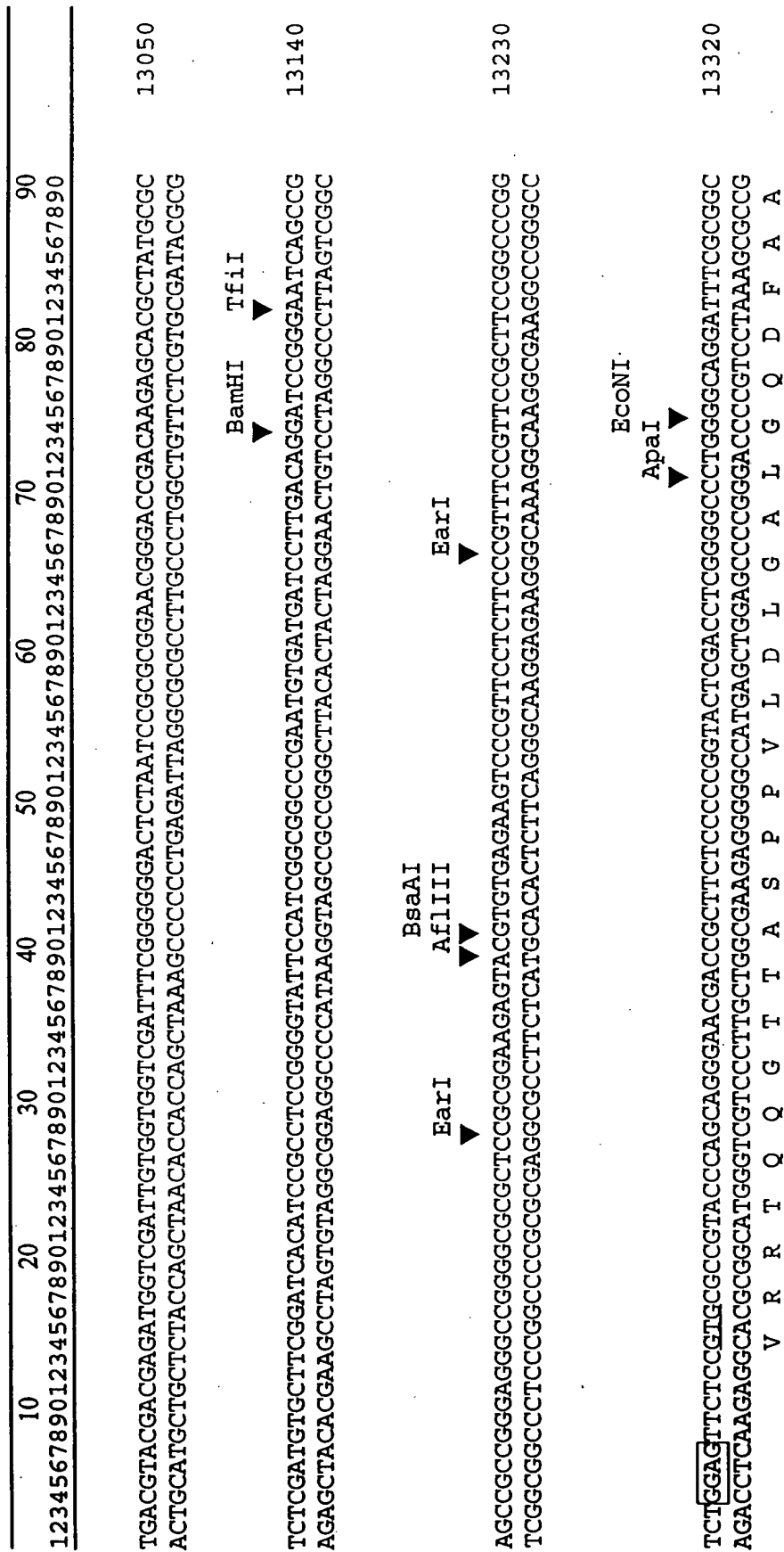


FIG. 32 - 37

sugar.finalgene b-1 Sequence

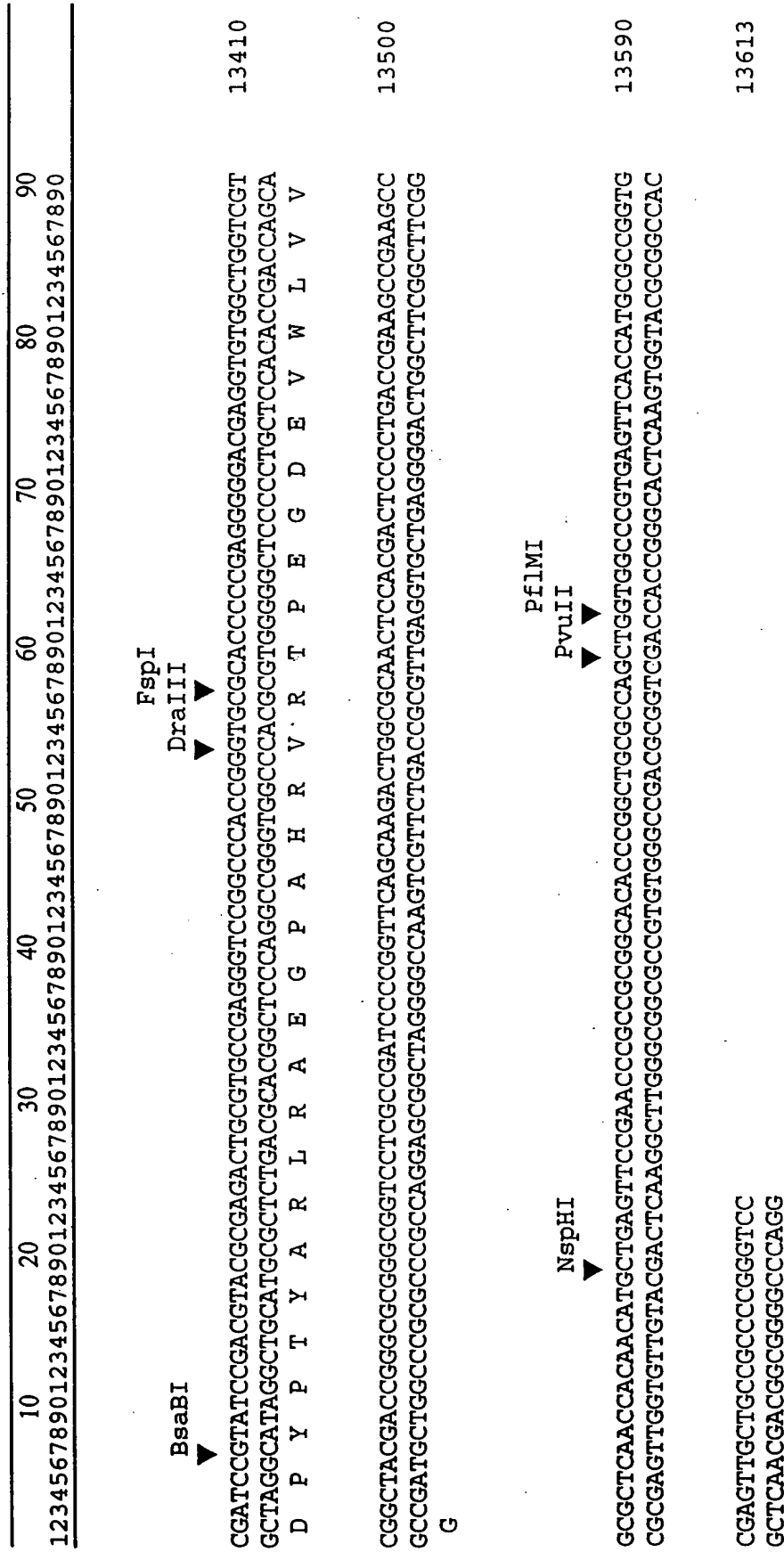


FIG. 32 - 38

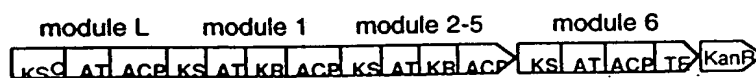
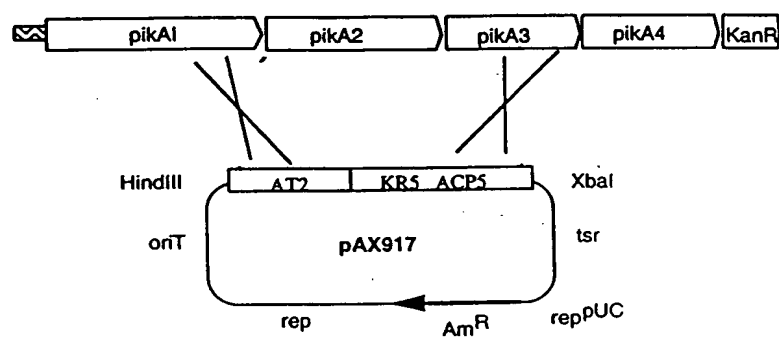
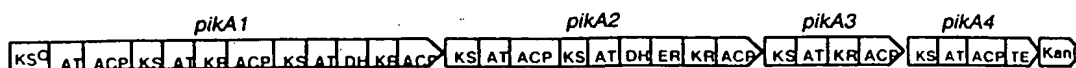


FIG. 33

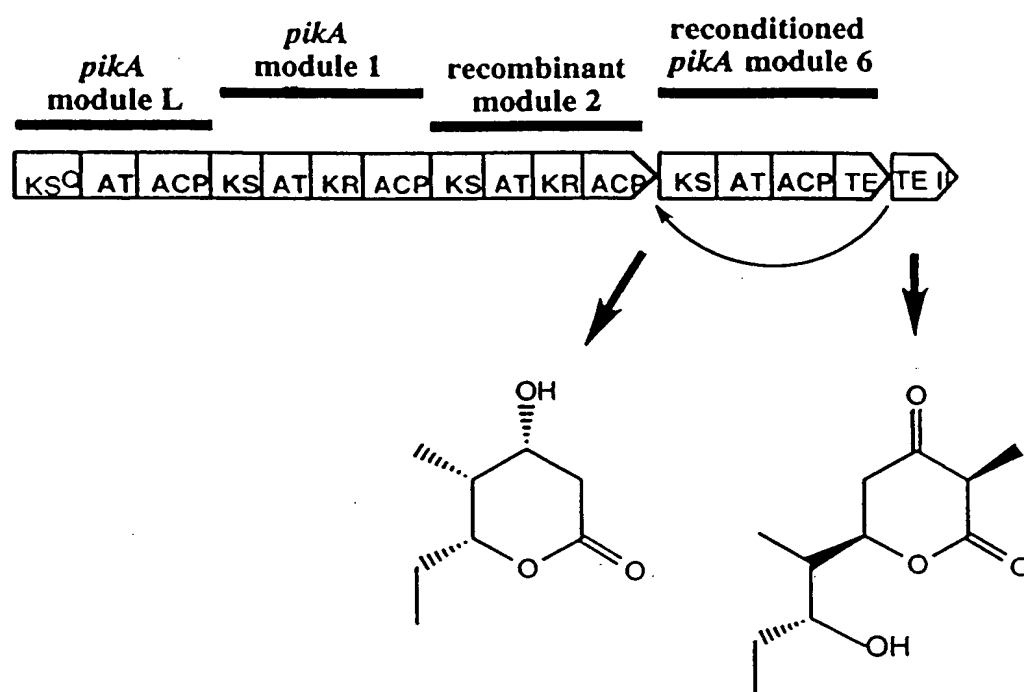


FIG. 34

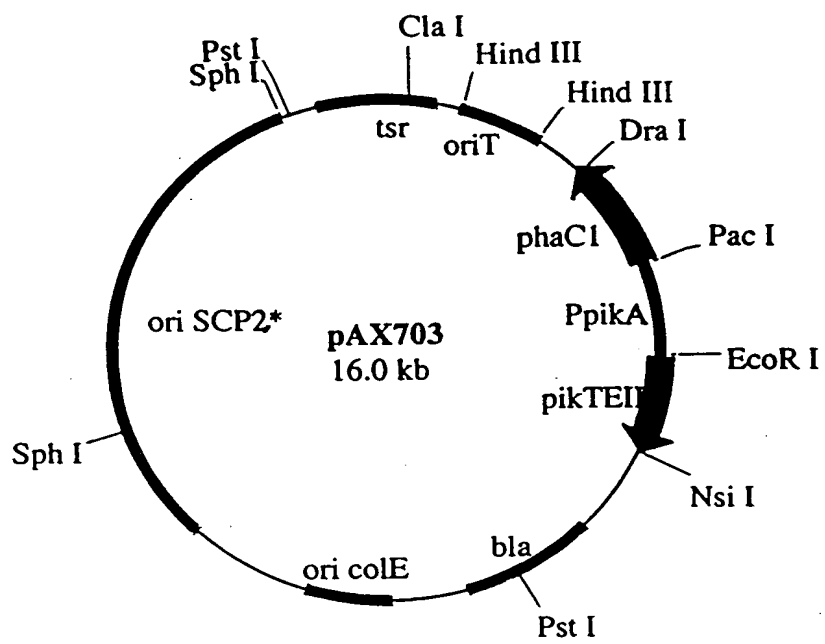


FIG. 35

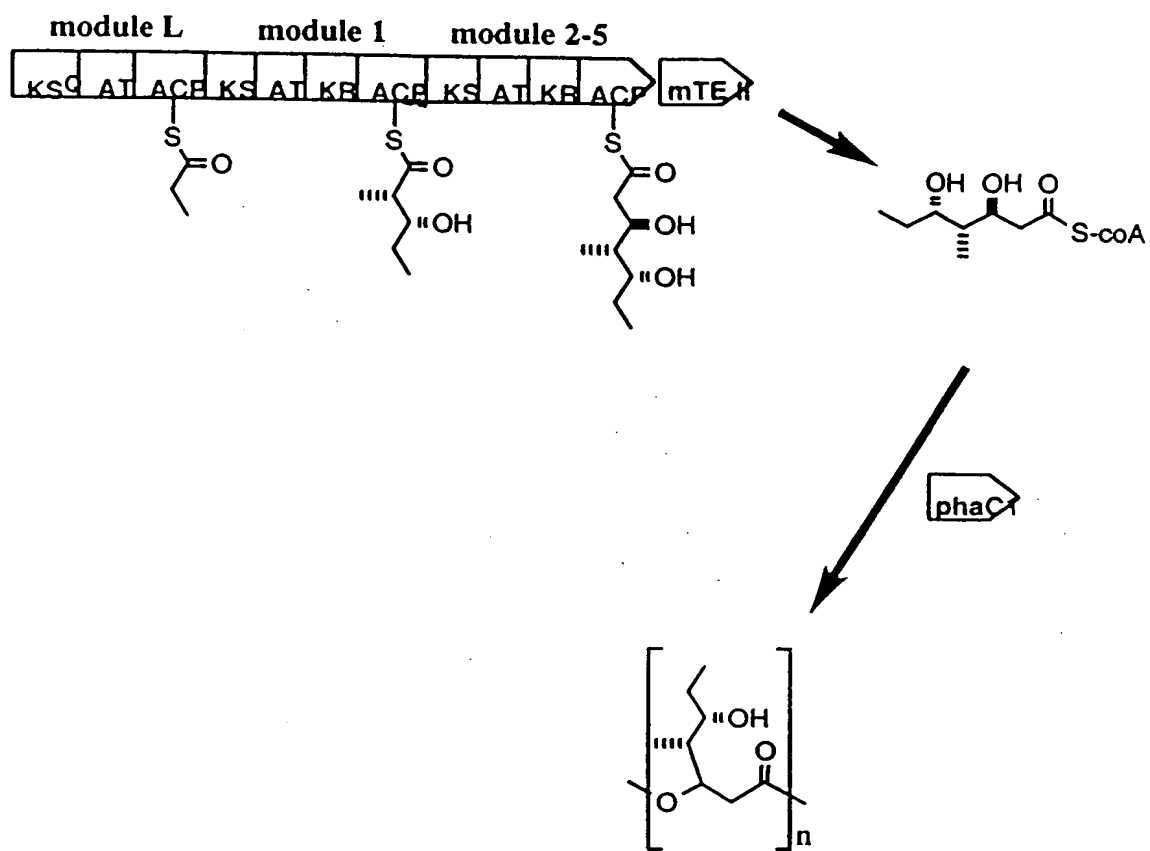


FIG. 36

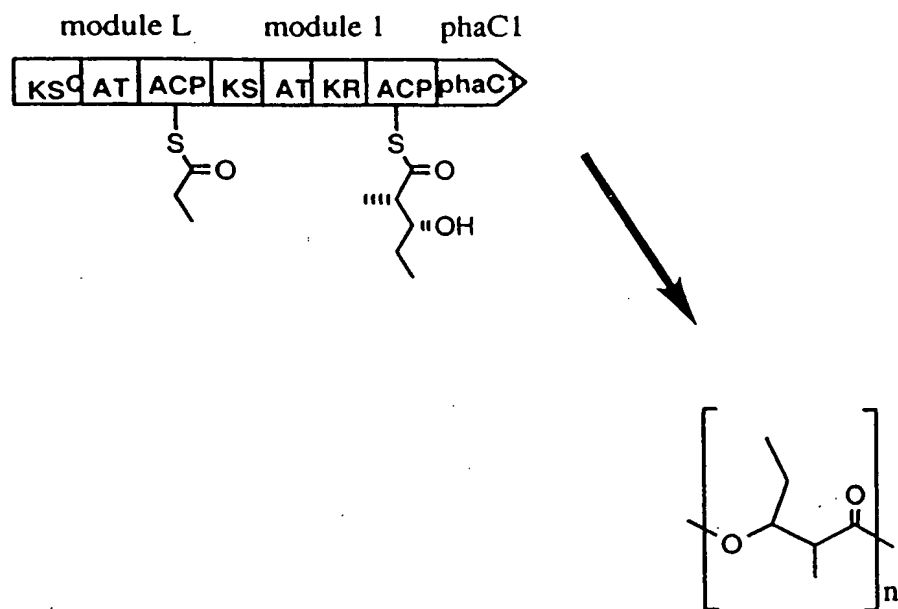


FIG. 37

<u>Amino Acid</u>	<u>Codon</u>
Phe	UUU, UUC
Ser	UCU, UCC, UCA, UCG, AGU, AGC
Tyr	UAU, UAC
Cys	UGU, UGC
Leu	UUA, UUG, CUU, CUC, CUA, CUG
Trp	UGG
Pro	CCU, CCC, CCA, CCG
His	CAU, CAC
Arg	CGU, CGC, CGA, CGG, AGA, AGG
Gln	CAA, CAG
Ile	AUU, AUC, AUA
Thr	ACU, ACC, ACA, ACG
Asn	AAU, AAC
Lys	AAA, AAG
Met	AUG
Val	GUU, GUC, GUA, GUG
Ala	GCU, GCC, GCA, GCG
Asp	GAU, GAC
Gly	GGU, GGC, GGA, GGG
Glu	GAA, GAG

FIG. 38

Original Residue	Exemplary Substitutions	Preferred Substitutions
Ala (A)	val; leu; ile	val
Arg (R)	lys; gln; asn	lys
Asn (N)	gln; his; lys; arg	gln
Asp (D)	glu	glu
Cys (C)	ser	ser
Gln (Q)	asn	asn
Glu (E)	asp	asp
Gly (G)	pro	pro
His (H)	asn; gln; lys; arg	arg
Ile (I)	leu; val; met; ala; phe norleucine	leu
Leu (L)	norleucine; ile; val; met; ala; phe	ile
Lys (K)	arg; gln; asn	arg
Met (M)	leu; phe; ile	leu
Phe (F)	leu; val; ile; ala	leu
Pro (P)	gly	gly
Ser (S)	thr	thr
Thr (T)	ser	ser
Trp (W)	tyr	tyr
Tyr (Y)	trp; phe; thr; ser	phe
Val (V)	ile; leu; met; phe; ala; norleucine	leu

FIG. 39

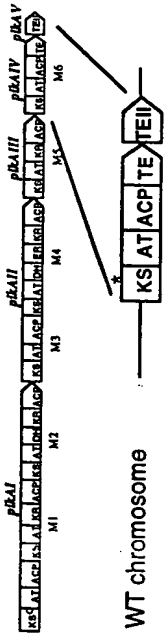




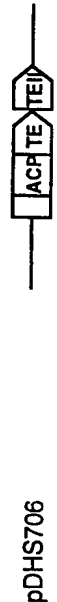


Genotype	10-Deoxymethynolide production (SCM medium)	Narbornolide production (PGM medium)
 <p>WT chromosome</p>	100%	100%
 <p>AX912 chromosome</p>	0	0
 <p>pDHS702</p>	trace	trace
 <p>pDHS704</p>	~60%	0
 <p>pDHS705</p>	~20%	trace
 <p>pDHS706</p>	~30%	~10% narbornolide trace 10-deoxymethynolide
 <p>pDHS708</p>	~100%	~50% 10-deoxymethynolide ~50% narbornolide
 <p>pDHS707</p>	~40% 10-deoxymethynolide ~60% narbornolide	100%

FIG. 40

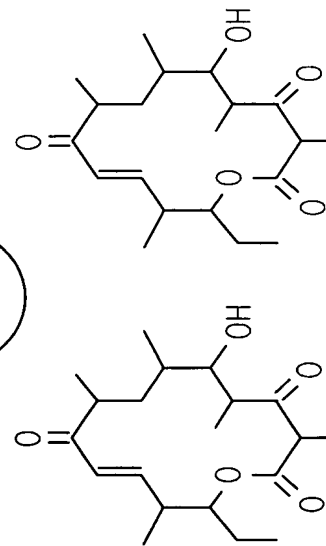
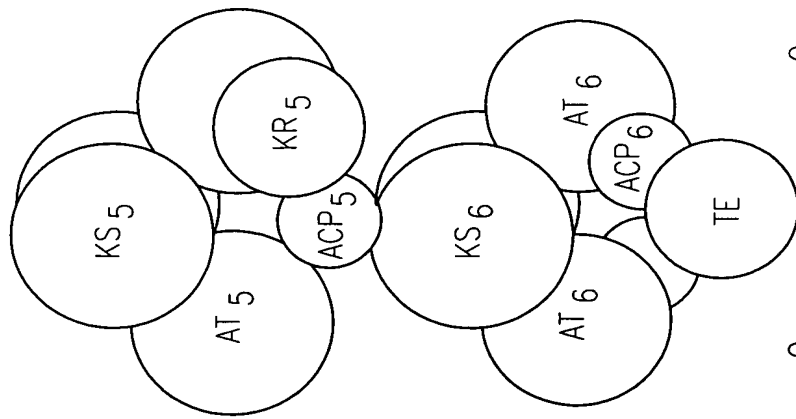


FIG. 41A

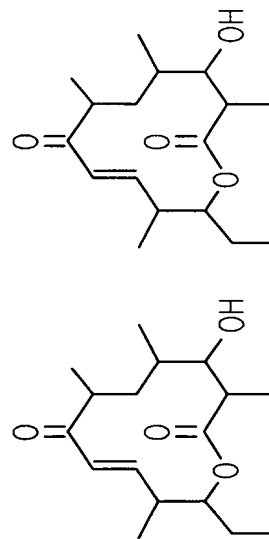
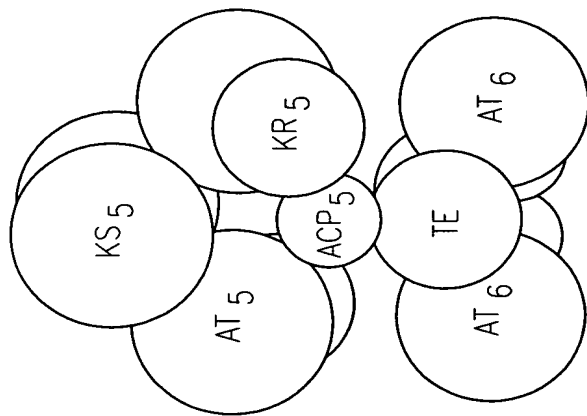


FIG. 41B

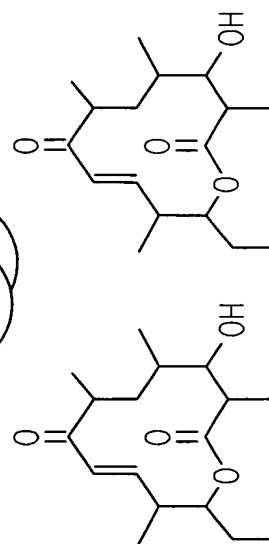
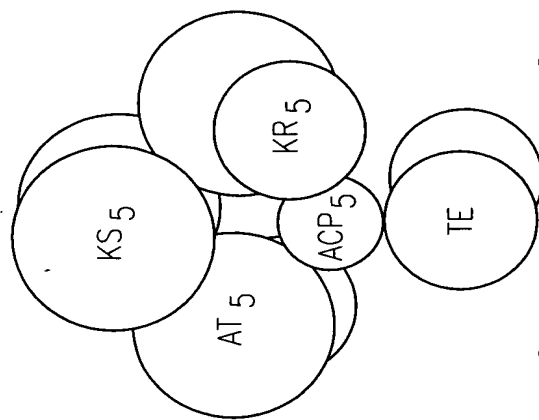
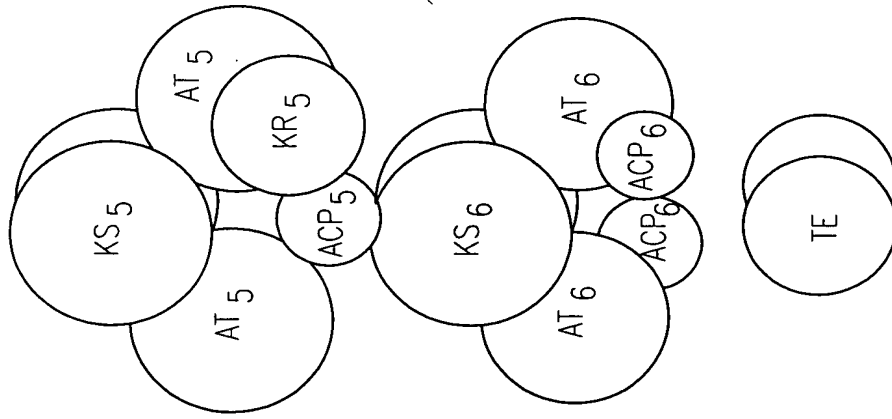


FIG. 41C



NO PRODUCT

FIG. 41D

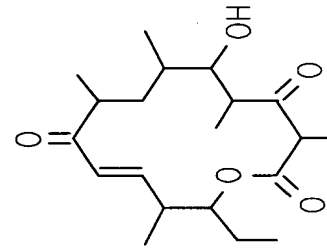
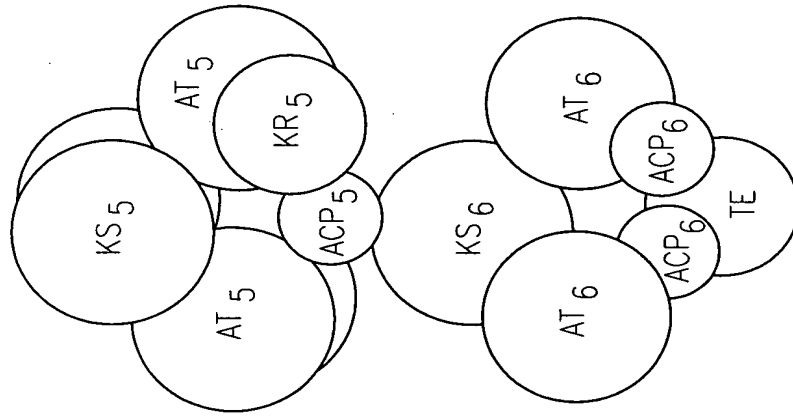


FIG. 41E

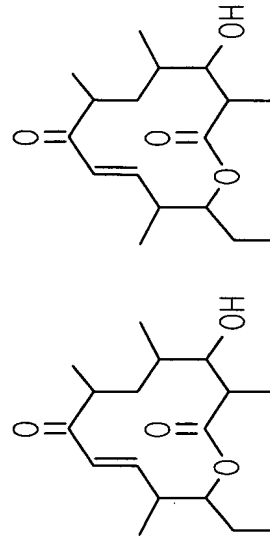
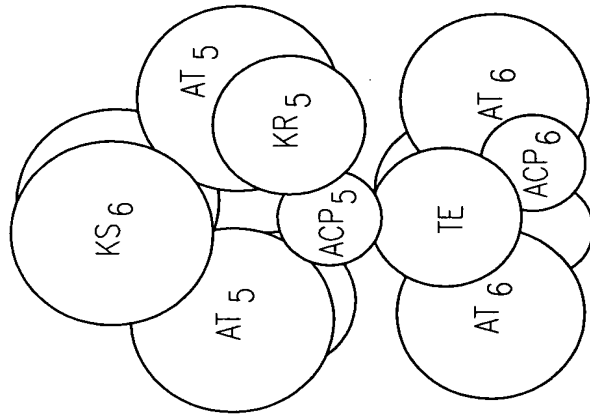


FIG. 41F

Scheme 1

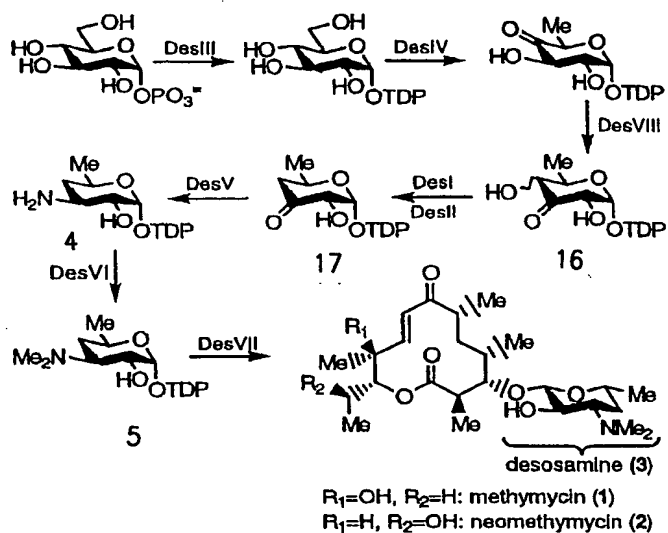


FIG. 42

Scheme 2

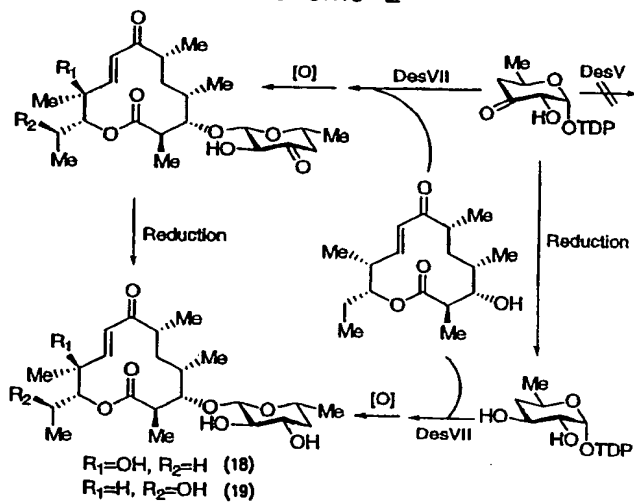


FIG. 43

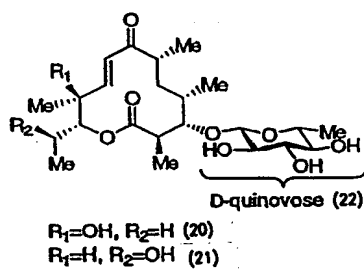


FIG. 44

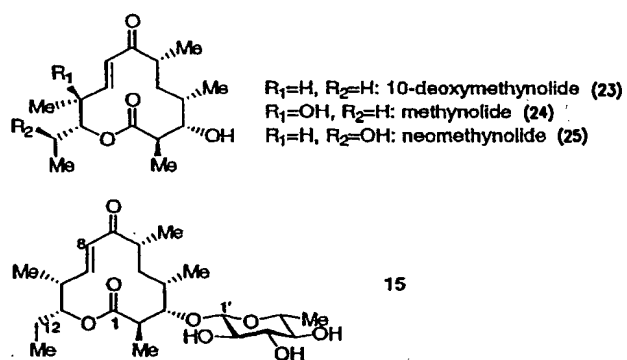


FIG. 45

A) . pika3-pika4 region from *Streptomyces venezuelae* ATCC15068

TGGCCGAAGTCCCCCTCGACCGGCTGCGGGACGCCGGGGTCTCTGACACCGTCTCTGCGGCTCACCGGAATCGAGCCCGAG
CCGGTGTCCGGCGGCCCGGACGGCGCCCGGACCCCGGTGCGGAGCCGGAGCCGGAGACGTCGATCGACGACCTCGA
CGCCGAGGCCCTGATCCGGATGGCTCTCGGCCCGCGGAACACCTGACCCGACCGCGGCCACGGCCCGCACACCGCCAGGT
GCCGTAGGCACACCCGACCCCTGCCCCCACAAGCCCAACATCCACGAGCGGAAGACCAACCCAGATGACG
AGTTCCAAACGAGCAGTTGGTGGACGCTCTGCGCGCTCCCTCAAGGAGAACGAAGAACTCCGGAAAGAGAGCCGTCGCCG
GGAC

B) . pika3-pika4 region from *Streptomyces narbonensis* ATCC19790

TGGCCGAAGTCCCCCTCGACCGGCTGCGGGACGCCGGGGTCTCTGACACCGTCTCTGCGACTCACCGGCATCGAGCCCGAG
CCGGTGTCCGGCGGCCCGGACGCTCGCCCGGCCCGCCCGCGGATCCGGAACCGGAGACGTCGATCGACGACCTCGA
CGCCGAGGCCCTGATCCGGATGGCTCTCGGCCCGCGGAACGCCCTGAGCACCCGCCCGGCCCGTGGTGTGCCCGGCCCTT
GCCGACTGCGGGCGGGCCCGGGCCCGCACACCGCCACGTACCAACCCGACACCGCCCGCCACACGCCCCACAACGC
CATCCACGAGCGGAAGACCAACCCAGATGACGAGTTCCAAACGAGCAGTTGGTGGACGCTCTGCGCGCTCTCCCTCAAGGA
GAACGAAGAACTCCGGAAAGAGAGCCGTCGCCGGGAC

C) . TE II gene from *S. venezuelae* ATCC15068

TCGACGGCATCGAGCGGGACACCGCCCGGGACCGCGGGACCGCGGAGGGGGGGAACAGTGACCGACAGCGTTCTGAACG
TGACGGCAACCTGTGGATCCGGCGCTTCCATCCCTCGCCGAACAGCGGGTGGGACTGGTCTGCCCTGCCCATGCCCCGC
GGTCCGCCAGTACTTCTCCGCTTCTCGGAGGAGTCCACCCCTCGTGGAGGCCCTGTGCGTGCAATACCCGGGCCG
CCAGGACCGGCGTCCGAGCCGTGCCCTGGAGAGCGTCGAGGAGCTGGCCGAGCAGTGGTCCGGGCCACCCGACCCCTGGT
GGCAGGA

D) . TE II gene from *S. narbonensis* ATCC19790

TCGACGGCATCGAGCGGGACACCGCCCGCGGGCGGACCGCGGAGGGGGCGGACAGTGACCGACAGAGTTCTGAACG
TGACAGCAGCCTGTGGATCCGACGCTTCCACCCCTCGCCGAACAGCGGGTGGGCTGGTCTGTGCGCGACGCCGGT
GGTTCGCCAGTACTTCTCCGCTTCTCGGAGGAGTCCACCCCTCGGTGAGGCCCTGTGCGTGCAATACCCGGGCCG
CCAGGACCGGCGTCCGAGCCGTGTCTGGAGAACGTCGAGGAGCTGCCGAGCAGTGGTCCGGGCCACCGAAGCCTGGT
GGCGGGA